国外优秀教材

PRINCIPLES OF
MEDICAL GENETICS

医学遗传学原理

第二版

T. D. 盖莱哈特
〔美〕F. S. 柯林斯 著
D. 金斯伯格

孙开来 等 译

科学出版社

(R-0586.0101)
责任编辑：马学海　谢灵玲　封面制作：王　浩　责任校对：白　颖

P R I N C I P L E S O F
M E D I C A L G E N E T I C S

医学遗传学原理

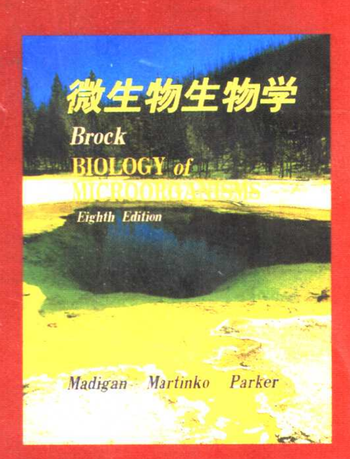


第二版



书　名　医学分子生物学（影印版）　　2000年6月出版
Molecular Biology in Medicine　　定 价 55.00 元
作　者　Timothy M. Cox & John Sinclair
书　号　7-03-008503-5/R · 540

本书由剑桥、牛津等名校的医学院资深教授和学者联合编著而成。内容涉及基因结构、基因表达、基因及药物治疗、遗传缺陷与人类疾病，包括微生物感染、病毒感染、血友病、免疫系统紊乱及癌症等，是一本有关分子生物学概念、原理和方法在医学中应用的指导性书籍。视角独特，思路清晰。



书　名　微生物生物学（翻译版）　　2001年3月出版
作　者　〔美〕M. T. 马迪根　J. M. 马丁克　J. 帕克　　估 价 115.00元
译　者　杨文博 等
书　号　7-03-008345-8/R · 954

*Brock Biology of Microorganisms*是国外的经典教科书。本书系其第八版的中文翻译版本，包括23章、附录及词汇表。该书从细胞化学和细胞生物学入手，详尽地介绍了微生物的营养、代谢、合成；分子遗传和微生物遗传；遗传工程和生物工程；生长和调控；工业微生物；寄主和寄生菌的关系；免疫学和免疫；临床和诊断微生物；流行病学和公共卫生微生物学；主要的微生物疾病；微生物代谢的多样性；微生物生态学；分子体系和微生物进化；病毒、细菌、古细菌和真核微生物等内容。该书阐述清晰，条理性强，简明易懂，并附有图表，每节后有概括的内容总结，便于读者学习。



科学出版社 生命科学编辑部
http://spbio.yeah.net
e-mail:spbio@163.net

ISBN 7-03-008746-1/R · 586
定 价：42.00 元

医学遗传学原理

〔美〕T.D. 盖莱哈特　F.S. 柯林斯　D. 金斯伯格　著
孙开来 等　译

科 学 出 版 社
2001

内 容 简 介

近年来人类分子细胞遗传学迅猛发展，尤其在人类基因组计划方面取得了重大成就。本书运用这些新成就，结合遗传学基本理论，系统而有重点地阐述了遗传学在医学中的重要作用；基因和染色体的结构与功能；群体遗传学与多基因遗传；现代分子生物学各种技术；人类疾病的细胞与分子遗传学机理；致病基因的定位与克隆；人类基因组学；临床遗传学；癌遗传学与基因治疗等章节。书后还附有每章的问题与答案，有助于加深对内容的理解。

本书是世界著名的遗传学家，美国 NIH 人类基因组研究所主任 F.S. Collins 和密歇根大学 T.D. Gelehrter 教授等编著的医学遗传学名著《医学遗传学原理》的第二版，译文较好地保持了原书的风格，是一本值得推荐的教科书和参考用书。适合作为医科院校的本科教材，也可供临床医生和医学遗传学研究人员及相关研究人员参考使用。

This is a translation of

Principles of Medical Genetics, Second Edition

Thomas D. Gelehrter, Francis S. Collins, David Ginsberg

* 本书中提供了一些药物使用的明确指示、不良反应、剂量和给药时间等，但这些都有可能发生变化。强烈建议读者仔细阅读药品包装上由其制造商提供的数据和治疗说明。

图书在版编目(CIP)数据

医学遗传学原理/(美)盖莱哈特(Gelehrter, T.D.)(美)柯林斯(Collins, F.S.),(美)金斯伯格(Ginsburg,D.)著；孙开来等译. -北京：科学出版社，2001.2

ISBN 7-03-008746-1

Ⅰ.医… Ⅱ.①盖…②柯…③金…④孙… Ⅲ.医学遗传学 Ⅳ.R394

中国版本图书馆 CIP 数据核字(2000)第 68301 号

科学出版社 出版

北京东黄城根北街16号

邮政编码：100717

新蕾印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2001年2月第 一 版 开本：787×1092 1/16

2001年2月第一次印刷 印张：25 插页：2

印数：1—5 000 字数：568 000

定价：42.00元

（如有印装质量问题，我社负责调换〈环伟〉）

译 者 名 单

主　译：孙开来

副主译：熊第志　赵彦艳

译　者（依姓氏笔画为序）：

马鸣超　王　敏　王红杨
付　浩　吕晶玉　孙开来
孙秀菊　孙　淼　李福才
邱广蓉　邱广斌　房　凯
赵　震　赵彦艳　贺　光
康　宁　曹　慧　富伟能
熊第志

中译本序

很高兴也很荣幸应邀为《医学遗传学原理》中译本作序。谨代表我的合作者 Thomas Gelehrter 和 David Ginsburg,祝贺孙博士和他的同事们完成了这项翻译工作,使中国的医学生和卫生保健专业人员更易理解遗传医学这个迅速发展领域的原理。仅在短短的几年中,遗传学从一门神秘的学科发展成为生物医学研究中主要的科学。在未来的几十年中,遗传学在医学的主流实践中将显示突出的作用。因此,每位医生必须熟知这个领域的原理,并准备将其应用于医治繁多病种的患者。

作为国际人类基因组计划的结果,遗传学在过去的十年中有着引人注目的快速进展。2000 年 6 月宣布了覆盖人类 DNA 90%的"工作草图"已经完成。更值得高兴的是,这项工作是世界各地 16 个中心共同完成的。他们孜孜不倦地工作,极其快速地取得大量数据,并在每天将这些信息输送到公共区域。为此作出历史性里程碑贡献的 6 个国家之中,就有中国。因此,您的国家应当分享这一重大成就所带来的荣誉。

为了大家的利益,我希望中译本的读者能够理解遗传学变革在所有医学领域的巨大潜能,并将有助于在中国使这些可能前景变为现实。

Francis S. Collins
医学博士,理学博士
美国卫生研究院人类基因组研究所所长

译 者 序

《医学遗传学原理》一书的中译本出版了。这是我国医学遗传学教育的一个重要事情。此书自1991年问世后,我们就采用为我校七年制和英语医学班的教材。多年来的教学实践,我们体会到这是一本很好的医学遗传学教科书。我们也曾向国内医学院校推荐此书,许多单位也采用为教科书或参考书。作者对遗传学的基本理论知识紧密结合医学实践作了深入而精确的论述。通过疾病来阐述遗传学的基本原理是此书的最主要特色。1998年的第二版又补充了近年来细胞和分子遗传学进展的新成果,特别是人类基因组计划取得巨大的成就,使医学遗传学已成为医学中发展最快的学科,也是现代医生和医务人员必修之课。正如 Collins 在为中译本所写序言中指出:“在未来的几十年中,遗传学在医学的主流实践中将显示突出的作用。”

中译本的出版得到科学出版社生物编辑部的大力支持。特别是原著作者之一,世界著名遗传学家,美国国立卫生研究院国家人类基因组研究所主任 F.S.Collins 博士代表三位作者为中译本撰写了序言,这无疑是对我们也是对我国医学遗传学界的巨大鼓励。中译本的问世,也是中国医科大学医学遗传学教研室全体师生共同努力学习的成果。由于我们水平有限,译文难免有错误之处,敬请医学遗传学界的专家和同志们予以指正。

我们希望《医学遗传学原理》一书中译本的出版将对我国医学遗传学教育的普及与提高起到促进作用。

孙开来

2000年10月

目 录

第一章　遗传学在医学中的作用

大自然运行的轨迹万一有显露她脱离常轨之处，那正是她更易于展露其神秘之处；欲要推进医学朝着正确的实践方向发展，最好的方法也莫过于潜心通过仔细研究稀罕病例来揭示普遍的自然规律了。因为，人们业已发现，在几乎所有的事物中，它们所蕴涵的普遍适用的特性，除非我们失去了拥有它们的权利，或者它们在某些方面自乱阵脚，否则我们是很难察觉这些特性的。

——William Harvey，1657

人们对遗传因素在人类疾病发生中的作用的认知，促使临床遗传学成为医学中发展最迅速的学科之一。随着在发达国家营养性和传染性疾病的明显减少，人们对人类疾病的遗传因素的决定作用愈加重视；业已证明，重大疾病如冠状动脉疾病、糖尿病、高血压和主要精神病的病因具有重要遗传因素。与此同时，分子遗传学和基因定位的进展促进了基础遗传学理论知识的爆炸性发展。已鉴定大约600个染色体的基因座，在每个位点上都有一个或多个特异的致病性突变；许多这些新资料的信息资源被直接应用于深入阐明疾病的发病机制和完善病人的诊断和治疗。更恰当地讲，遗传学这些新进展的重大贡献就在于预防和(或)防止疾病，这一方向必然是现代医学的焦点。检测个体发病风险的遗传筛选、完善的遗传诊断、遗传咨询和产前诊断都是新的遗传学知识在现代医学实践的应用。基因治疗试验也已经开始应用于治疗特异性疾病，并将对未来的临床医学产生巨大影响。最近，医学遗传学已成为美国医学专业委员会的第24个成员，这是在12年内所认识的第一个新专业。鉴于遗传学对预防和防止临床疾病的贡献越来越大，遗传服务一定能成为任何一个新的健康保险计划的一个组成部分。

一、遗传病的影响

与普通的观念相反，许多遗传病并非罕见，而且实际上是多种疾病和死亡的重要原因。遗传病从个体上看是稀少的，但从整体上看却是发病率和死亡率的一个主要原因。大约有3%的孕产婴儿患有明显遗传病，或者有跛足、智力低下等出生缺陷，或者夭折。在不列颠哥伦比亚省，对一百多万25岁以下个体的调查显示至少1/20个体患有由重要遗传因素造成的严重疾病。许多遗传病的慢性特征给病人、家庭和整个社会带来了沉重的医疗、经济和精神负担。

在英国，对1 200多名住院儿童死因的两项调查显示，遗传性疾病占死亡率的38%～42%(表1-1)。在北美的两项调查发现，近17 000名住院儿童中明显的遗传性疾病占

表 1-1　英联邦中幼儿死亡中的遗传成分*

死亡原因	纽卡斯尔	伦　敦
染色体的	2.5%	12.0%
单基因的	8.5%	
多基因的	31.0%	25.5%
非遗传的/未知的	58.0%	62.5%
总死亡数	1041	200

* 引自1. Robert DF, Chavez J, Court SDM: The genetic component in child mortality. Arch Dis Child 45: 33 ~ 38, 1970.
2. Carter CO: Changing patterns in the causes of death at the Hospital for Sick Children. Great Ormond St J 11: 65 ~ 68, 1956.

5% ~ 10%。如果把与遗传因素相关的疾病包括在内,则有三分之一到一半以上的患儿为遗传性疾病(表 1-2)。而且,遗传病患者住院次数和时间均增加。虽然在成人中由遗传引发的疾病发病率还不清楚,但估计至少 10%的住院成年病人为遗传性疾病。因此,遗传性疾病确实造成了严重的经济和医疗负担。

表 1-2　北美医院儿科就诊患者中遗传病所占有的百分率*

原　因	西雅图	蒙特利尔
染色体的	0.6%	0.4%
单基因的	3.9%	6.9%
多基因的	48.9%	29.0%
非遗传的	46.6%	63.7%
就诊总数	4 115	12 801

* 引自1. Hall JG, Powers EK, Mcllvaine RT, Ean VH: The frequency and financial burden of genetic disease in a pediatric hospital. Am J Med Genet 1:417 ~ 436, 1978.
2. Scriver CR, Neal JL, Saginur R, Clow A: The frequency of genetic diseases and congenital malformation among patients in a pediatric hospital. Can Med Assoc J 108: 1111 ~ 1115,1973.

除了外伤,“非遗传性”一词就成了误称,因为难以想像任一疾病是完全非遗传的。任何个体的发育取决于遗传和环境因素的相互作用。在整个发育过程中,遗传因子尽管表达各异,但从妊娠开始就一直存在,而环境影响在不断变化。由于所有人体健康和疾病的改变在某种程度上都具遗传性,因此所有疾病都是遗传性的。曾经有人认为传染性疾病是非遗传性疾病的鲜明例证,因为已经确定了该病的特异外源性病原体。然而,现已知道许多宿主防御因子是由遗传决定的,它们在感染易感性以及对传染源的免疫应答方面起着十分重要作用。因此,即使是有明确的外源性病因的疾病,遗传因素仍起关键作用。这种理论可更精妙地解释像酒精中毒这样常见问题的原因。

遗传因素在人类变异和疾病中的作用特点和作用程度是人类遗传学和医学遗传学领域的研究内容。鉴定疾病的易感遗传因子以及具有遗传易感素质的个体是发现疾病重要环境因素的有力手段。

但看起来有些矛盾的是,鉴定疾病的遗传因子最重要的益处之一并非是其具有基因治疗的潜力(尽管那可以成为一个激动人心的希望),但更多的是对遗传易感个体采取环境监控,把握治疗和预防机会。

二、遗传病的主要类型

遗传性疾病通常分为三大类:染色体病、单基因病和多基因或多因素病。人类癌症分子基础研究的最新进展需要增加第四类,即体细胞遗传缺陷。每一类将在下面简单地讨论并在以后的每一章节更深入地详述。

(一) 染色体病

这类疾病是由于整条染色体或部分染色体的增加或缺失所致,将在第八章详细论述。因为每条染色体有数万个基因,所以染色体病的体征通常十分明显,大多数染色体病的特征为生长迟缓、智力低下和各种身体异常。临床上典型的染色体异常约占初生儿的 1%,住院儿童的 1%,死亡儿童的 2.5%。失去或获得整条染色体一般难以存活,是自然流产或夭折的主要原因。染色体异常几乎占自然流产的一半,约 15%的妊娠已知以流产告终,这样看来 1/4 的胎儿可能出现染色体异常。因此,染色体疾病的主要影响发生在出生前(图 1-1)。

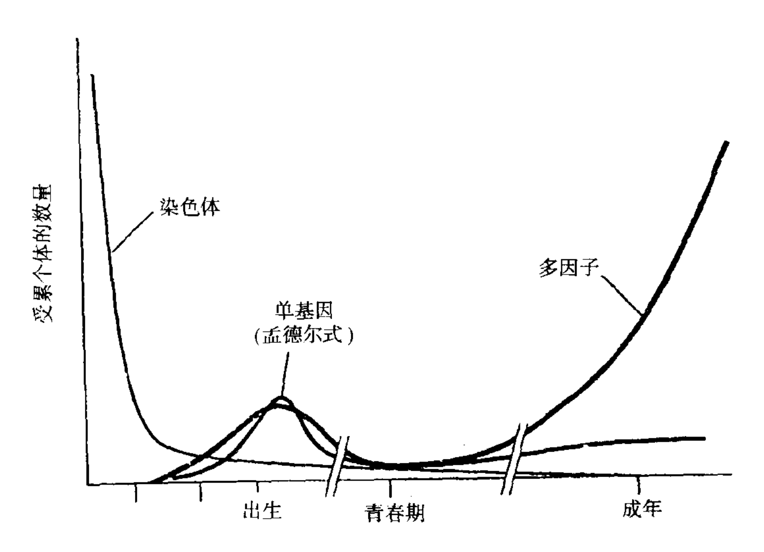


图 1-1 主要类型遗传疾病的表达年龄。该模式图显示了患有染色体疾病、单基因疾病和多因素遗传疾病的患者产前和出生后的不同时段的相对数量。

主要染色体疾病的典型例子是 Down 综合征,由 21 三体型所致,即 21 号染色体有三个拷贝,而不是通常的两个。这种疾病约占活婴的 1/800,且随着孕妇年龄增长而增加。Down 综合征表现为生长迟缓,不同程度的但多为严重的智力低下以及身体异常,包括上斜眼,过去曾不幸地称为"蒙古症"。其中与之相关的大多数先天畸形为先天性心脏缺陷,这是 Down 综合征儿童死亡的主要原因。21 三体型也明显降低了胎儿的宫内存活率,而受累胎儿多数自然流产。Down 综合征是在人类中发现的第一个染色体病,并且首次表明染色体物质的改变可引起智力低下和严重的先天异常。Down 综合征也是最先进行羊

水穿刺产前诊断的疾病之一。

(二) 单基因病

单基因病是由单个基因突变所致,严重影响患者的健康。正如所料,单基因病是按单纯的孟德尔方式遗传(见第三章),所以又称为孟德尔疾病。现在约有 4 000 多个不同的疾病确定为单基因病,以常染色体显性、常染色体隐性或 X 连锁方式遗传。其中某些疾病将在以后的章节详细论述。单基因病占儿科住院人数和儿童死亡率的 5% ~ 10%。尽管单基因病在成人中的重要性正在逐渐受到重视,但它主要发生在新生儿和幼儿阶段(图 1-1)。虽然许多单基因病比较罕见,但有些却很普遍并危害健康。家族性高胆固醇血症有较高的早期冠状动脉疾病发病风险,发病率约 1/500。家族性乳腺癌和遗传性结肠癌的发病率约为 1/300。在美国每 400 个黑人中有 1 人患镰状细胞贫血,每 2 000 个白人中有 1 人患囊性纤维化。镰状细胞贫血是第一个在分子水平上确定的遗传性疾病,它的研究成为现代分子遗传学应用于临床疾病的一个模型。这足以证明单基因病是分子遗传学进展对了解和治疗疾病作出重大贡献的领域。

线粒体染色体上的基因同核染色体上基因一样也能发生突变。线粒体疾病通常影响神经和肌肉的能量产生,在细胞衰老中起作用。这些疾病以一种独特的母系方式遗传,将在第三章进行讨论。

(三) 多基因或多因素疾病

这类疾病是由多个基因相互作用所致,其中一些可能起主要作用,但多数为相对微效作用。这类疾病既是最常见的又是最不了解的人类遗传性疾病。掌握常见慢性疾病的遗传学基础代表了当代医学遗传学所面临的挑战。这种多基因疾病包括成年的常见疾病如糖尿病、高血压、冠状动脉疾病和精神分裂症,还包括常见的先天缺陷如唇裂、腭裂和许多先天性心脏病。这些疾病占儿科住院的 25% ~ 50%,约占儿童病死率的25% ~ 35%。许多这类疾病的长期性可能对成人造成沉重负担。因此,多基因病在新生儿和成人时期的临床影响都是非常重大的(图 1-1)。从理论上说,探索几个或众多基因间累加效应或相互作用而导致疾病易感性以及适当环境因素的影响,具有一定挑战性。希望分子遗传学方法、基因定位和遗传流行病的结合将有助于疾病特异性的遗传因子和决定疾病易感性的遗传异质性的鉴定。相互作用如何引起疾病的模式,确定遗传因子性质及其在这些疾病中作用的方法将在第四章和第九章中讨论。

(四) 体细胞遗传病

体细胞遗传病同上述三大类疾病不同点在于三大类疾病的遗传异常是在人体所有细胞包括生殖细胞(精子和卵子)的 DNA 中并能传递给下一代,而体细胞遗传疾病只在特异的体细胞中发生。体细胞遗传病的一个范例是癌,其恶性表型的发展通常是控制细胞

生长的基因发生突变所致。现已清楚,所有人类癌症都是 DNA 突变所致,使其成为最常见的遗传病。导致癌的各种遗传机制将在第十一章讨论。

三、遗传学与医学的相互渗透

几十年来,遗传学基础学科和医学临床学科一直是双向相互作用并且成果累累。在其早期阶段,遗传学倾向于植物和动物饲养,属昆虫学家的范畴,而并非医生的领域。孟德尔在 19 世纪 60 年代通过豌豆实验已能系统地表达基因概念。他采取计划交配,并观察数代,发现了决定性因素,而这在人类研究中是难以获得的。同样,Thomas Hunt Morgan 和其他人利用果蝇的传代期短和基因组相对简单(同哺乳动物相比)的特点,对果蝇(*Drosophila melanogaster*)进行研究,获益很大。然而,人类遗传学萌发于 20 世纪初,对一些重要理论的认识和探索要比其他物种更深奥。群体遗传学、多态性的研究(第九章)和生化遗传学(第七章)就是突出的例证。近几年,这种趋势仍在继续,新学科层出不穷。正如第十一章要论述的那样,人类肿瘤遗传学的新领域表明人类细胞携带一组基因叫"癌基因",另一组叫"肿瘤抑制基因"。它们在正常情况下参与形成失控的生长方式。这些基因的某些突变能使生长方式失控,我们认为是癌症。癌基因和肿瘤抑制基因的发现已证明了在癌症、病毒学和遗传学之间的一种不曾预料的联系,并把生物学和生化学的不同领域联系起来。许多癌症同特定的染色体重排有关,这也是重要的诊断和预后指标。克隆这些断裂点区域可以发现一些新的基因,并对正常和异常生长调控机制有更好的了解。而且,已经鉴定的几个癌基因或肿瘤抑制基因的遗传性突变可能导致出生缺陷而不是癌!最后,在家族性癌综合征中起作用的基因的克隆和鉴定揭示了在细胞生长的调控过程中的一些鲜为人知的步骤。例如,神经纤维增生症 I 型基因的克隆(第三章和第九章中讨论)得出一个惊人的发现:这个基因不只是肿瘤抑制基因,而且它的产物神经纤维素在涉及癌基因 *RAS* 和鸟苷三磷酸酶的正常生化信号传导通路中起关键作用!

因此,医学的成就已促使遗传学和基础生物学的不断深化发展。在另一方面,遗传学研究对医学的贡献甚至更为明显。在当今科技医疗保健系统中,我们很容易忘记医学在其大部分历史进程中是一个描述性的学科。虽然描述方法对于决定疾病的自然病史和记述各种治疗策略的效果是很有价值的,但医学上极为重要的进展却通常来自于对更为基础的科学原理的阐明和随后应用于临床实践。医学遗传学在最基本水平,即基因本身,研究人类疾病。因此,遗传学上的发展自然地对临床医学具有深远影响,而且这些影响将持续增长。

一个首要例子是分子病的概念,它是由 Pauling 根据镰状细胞贫血第一次提出的。这种疾病在西非作为一种持续一生的,能引起贫血,骨骼、关节和腹部疼痛的疾病已经被认识几百年了,但其实际病因却一直不清,直到 1910 年芝加哥的一个心脏病学家 James Herrick 首次发现在受累个体中存在异常形态的红细胞。四十年后,Pauling 证明了镰状细胞贫血患者血红蛋白的电泳特性与正常血红蛋白不同,而且受累儿童的父母显示出既有正常的,又有异常的血红蛋白。1956 年 Ingram 利用肽链指纹技术显示出这种差异是由于血红蛋白 β 链第 6 位的谷氨酸被缬氨酸替代所致。随后,在核苷酸水平确认突变是 T 代替

A 的单个碱基置换，而且现在可根据 DNA 分析进行产前诊断(详见第六章)。这样一种严重的和临床上复杂的疾病，只是由 30 亿个核苷酸中的一个核苷酸改变所致，这样的事实在前几个世纪却不能凭借那些成为医学特征的描述性方法所预测出来。这一事实有力地证明了把遗传学方法应用到医学上是极有意义的。它还表明了一个重要事实：单个基因变化能对多个器官系统具有复杂的临床效应。医学遗传学因此是一个宽广的学科，与其他所有临床学科有很大范围的重叠。

血红蛋白的例子还证明了第二个事实：突变的研究能对结构功能关系产生更透彻深入的认识。随着对镰状血红蛋白的描述，我们现在认识到它的破坏性是由于有利于血红蛋白在细胞内多聚化造成的。接着，又有 300 多个其他突变血红蛋白被描述。其中有些可能位于血红蛋白多肽链的非关键部位，没有功能上的后果；另一些位于血红素结合部位附近，能导致血红蛋白分子所携带的铁氧化(高铁血红蛋白症)。那些影响高(低)氧亲和力平衡的突变能产生高亲和力血红蛋白，不易释放氧。通过确定这些不同情况的突变位点，珠蛋白分子各个部位的功能作用开始显示出来。其他遗传病如葡萄糖-6-磷酸脱氢酶缺乏症、囊性纤维化、甲型和乙型血友病的详细分子水平分析结果证明这些疾病都类似地存在复杂的基因突变，可以改变或破坏相关基因产物的功能。

血红蛋白的研究对于我们了解人类生物学所作出的贡献难以言表，这也正是用这本教材的一整章(第六章)来讲述它的原因。它还揭示了另一个重要的道理：对一个特定的基因-蛋白系统进行仔细研究所学到的东西能使我们更好了解其他系统。

分子医学的概念不断地拓宽。例如，冠状动脉疾病是西方世界一个主要死亡原因。人们注意到高血清胆固醇与冠状动脉血栓形成风险增高的相关性已经有几十年了，但其遗传模式和导致胆固醇增高的代谢途径仍不清楚。具有高胆固醇水平的家族成员皮肤和肌腱内有胆固醇沉积，而且突发心脏病的危险性极高(见第七章)。Goldstein 和 Brown 集中对这类占心脏病 5% 的家族进行研究，证明细胞表面存在一个富含胆固醇的脂蛋白受体，称为低密度脂蛋白(low-density lipoprotein, LDL)，还发现这类患者存在 LDL 受体缺陷。用遗传方法阐明其重要的代谢途径对于我们了解胆固醇代谢和普通动脉硬化的原因有很大帮助。甚至不仅如此，它还扩展了我们对细胞表面受体生物学的整个现象的理解，包括这些受体是如何合成的，是如何定位在细胞膜上的，又是如何将它们的配体携带入细胞内部(这个过程称为受体介导的胞吞作用)。

这个例子还表明了另一重要方面，就是在本章开头引语中 William Harvey 清楚认识到的：对罕见遗传突变的研究可以对理解正常功能提供有力的途径。在现代医学中可以发现这种方法获得成功的许多例子。

四、小　结

几年前，仅几所医学院校开设医学遗传学课程。今天，大多医学生都得到医学遗传学的直接教授，而且这种重视有增加的趋势。医学遗传学与临床医学、医学遗传学与其提供的许多医学科学基础之间的直接关系使对这个领域的了解成为现代医生所必需的。幸运的是，掌握本书着重教授的一系列基本原理就可以应付复杂情况了。对于有许多东

西要记忆的医学生来说,遗传学可以成为一种极受欢迎的小憩——一个真正理性的、逻辑的、令人满意的学科。我们生活在一个遗传学发明的年代,同500年前的哥伦布和他当代的探险家和制图员一样令人激动。我们对我们人种遗传多样性的兴趣逐渐增加,使我们强调个体的特性。对疾病遗传基础的深入了解将对疾病预防起核心作用,这也是未来医学的焦点。

推 荐 读 物

Harvey W, quoted by Garrod A. The lesson of rare maladies. Lancet 1928; 1: 1055 ~ 1066.

McKusick VA. Medical Genetics. A 40-year perspective on the evolution of a medical specialty from a basic science. JAMA 1993; 270:2351 ~ 2356.

Motulsky AG. Michael Brown and Joseph Goldstein. The 1985 Nobel Prize in Physiology or Medicine. Science 1986; 231: 126 ~ 129.

Neel JV. Physician to the gene pool. Genetic lessons and other stories. New York: J Wiley & Sons, 1994.

Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus H. The molecular basis of blood disease. 2nd ed. Philadelphia: WB Saunders, 1994.

Weatherall DJ. The new genetics and clinical practice. 3rd ed. New York: Oxford Univ Press, 1991.

第二章　基因和染色体的结构和功能

明后天我和 Crick 将寄一篇短文给《自然》杂志，提出我们的(DNA)结构可能的模型，同时强调它的暂定特征和目前尚缺乏支持它的证据。即使错了，我认为也很重要，因为它提供由互补链组成一种结构的具体例子。如果这种模型正确，那么我想关于 DNA 以自我复制的方式，我们会做出些小贡献。

——James Watson，1953 年 3 月 12 日给 Max Delbrück 的信

为说明人类遗传学基本原理，我们必须从 DNA 开始。DNA 即脱氧核糖核酸，是生命的基本分子。精巧的人类基因组 DNA 编码着指导人类细胞生长、分裂、分化成特定的结构以及细胞对环境变化反应的所有指令。而且，所有遗传病的基础根本上都是建立在 DNA 序列的变化之上。总体上，对 DNA 及其组装形成的染色体的结构和功能的了解是医学遗传学和医学的基础，尤其当这些领域已经从描述性科学转向分子科学时。

本书的一些读者可能对本章及下章的内容已经很熟悉了。然而，本章之所以包含这些基本的内容，是因为考虑到对医学遗传学感兴趣的读者具有不同的背景。

一、DNA 是遗传物质

从 19 世纪 60 年代开始，人们就已经知道染色体含有 DNA 并引起遗传性状。但在 20 世纪较早年代，大多数生物学家都比较熟悉可能提供信息的蛋白质及其 20 个氨基酸的构件，因而倾向于认为染色体的蛋白质成分才是真正的遗传物质。直到由引起肺炎的肺炎球菌提供了一种遗传分子的生物学鉴定的发展才解决了这个问题。引起人类疾病的肺炎球菌是光滑的、发光的、有黏膜的菌落，而不致病的肺炎球菌外表粗糙(现在已经知道归因于是否存在或缺乏厚的多糖细胞壁)。活着的粗糙型细菌与经热处理杀死的光滑型菌落混合后，就能使一小部分转化为光滑型。1944 年，Avery、Macleod、MaCarty 等通过研究这种“转化因子”的性质确切地证明了这种“转化因子”的本质并不是蛋白质而是 DNA。这个基本的研究提供了一个新的基本范例并使遗传学进入现代革命的时期。

二、DNA 的结构

由于证明 DNA 本身携带遗传信息，在第二次世界大战末期，一些实验室把注意力转向 DNA 的物理结构。1953 年，Watson 和 Crick 应用 Wilkins 和 Franklin 精细的 X 射线晶体衍射分析数据推测出 DNA 的正确结构(图 2-1)。人们对 DNA 自身是一个长的多聚体并

不感到惊奇，但谁也没有预想到 DNA 是由相互逆向的两条链缠绕而成。

图 2-1 DNA 结构。左侧为 DNA 双螺旋的示意图，条带代表糖 - 磷酸骨架，碱基排列在中间。注意 A 总是与 T 配对，G 总是与 C 配对。右侧为一条链上 4 个核苷酸的扩展图示，表示 5′-ACGT-3′序列的全部化学结构（一个核苷酸由一个糖、一个磷酸及附着的一个碱基组成）。注意 5′和 3′的标志用来表示 DNA 链的极性，指的是一个脱氧核糖环上碳的序号。

DNA 的分子结构以图解描述如图 2-1 和图 2-2（空间模型），由双螺旋组成。双螺旋每条链的“骨架”由不变顺序糖 - 磷酸 - 糖 - 磷酸的多聚体构成，糖为脱氧核糖，磷酸通过酯键与其 5′和 3′羟基相连。糖环的 1′位与含有氮的四种“碱基”之一相连。其中两个是嘌呤碱基，即腺嘌呤（A）和鸟嘌呤（G）；另外两个是稍小一些的嘧啶碱基，即胞嘧啶（C）和胸腺嘧啶（T）。Watson 和 Crick 正确地推测出碱基几乎像一堆硬币一样，以每 3.4Å 相隔平铺于双螺旋内部。双螺旋结通过 A 和 T 碱基、G 和 C 碱基之间所形成的氢键相结合，其中每一个称为“碱基对”（base pair）（图 2-3）。因此 DNA 双链是“互补”的，若一条链为 5′-ATGCCAG-3′，那么另一条链一定是 5′-CTGGCAT-3′，这 7 个碱基对（bp）完全的双链结构

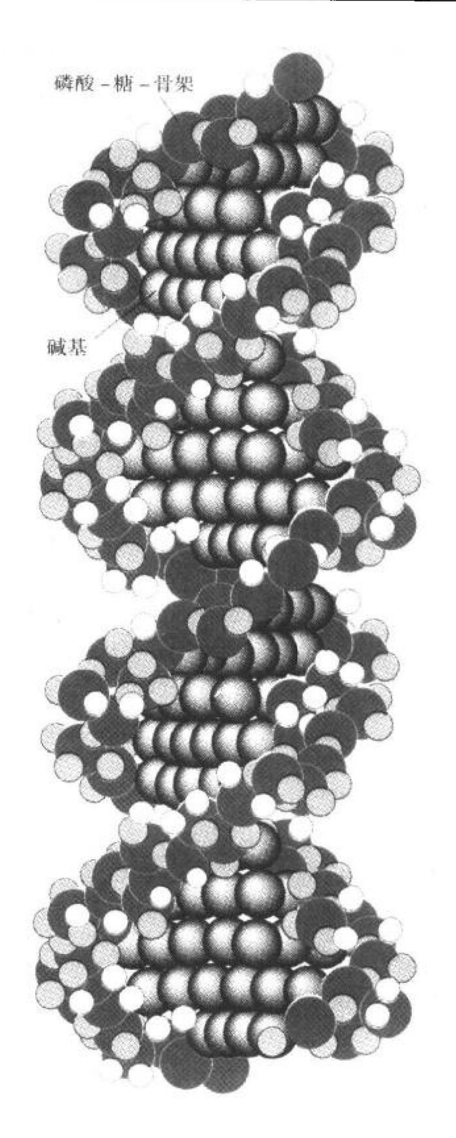


图 2-2　DNA 的空间模型。

如下：

5′-ATGCCAG-3′

3′-TACGGTC-5′

这种 A 与 T、G 与 C 的配对解释了 Chargaff 原则的基础，在 DNA 结构阐明以前就已推测出这一原则，该原则指出在一个特定物种的 DNA 中，A 和 T 碱基的百分率相同，G 和 C 的百分率也相同。

应该仔细地指明这种结构的若干特征：

1．根据 DNA 链中碱基的序列，它提供了储存和编码大量信息的方法。若一个分子长 N 个碱基，那么可能有 4^N 可能的排列序列。一个生物体的全部 DNA 序列含有它的全部遗传信息，称为基因组。最小病毒的基因组只有几千个碱基对，仅含有少数基因。然而，若沿着进化树前进，则基因组的大小及复杂性以非线性方式增加，从细菌的 4×10^6 bp

图 2-3　腺嘌呤 - 胸腺嘧啶碱基对和鸟嘌呤 - 胞嘧啶碱基对间的氢键。注意 A-T 对有两个氢键形成，G-C 对有三个氢键形成。

到人的 3×10^{9} bp。基因组大小的代表性例子如表 2-1 所示。

表 2-1　以碱基对代表基因组大小

SV40，一种哺乳动物 DNA 病毒	4×10^{3}
λ，一种细菌病毒	5×10^{4}
大肠杆菌（*Escherichia coli*），一种细菌	4×10^{6}
酿酒酵母（*Saccharomyces cerevisiae*），一种酵母	1.2×10^{7}
秀丽新小杆线虫（*Caenorhabitis elegans*），一种线虫	1×10^{8}
黑腹果蝇（*Drosophilia melanogaster*），一种果蝇	1.2×10^{8}
人类 21 号染色体	5×10^{7}
人类 1 号染色体	3×10^{8}
人类全基因组[a]	3×10^{9}

a. 这实际上是单倍体基因组，或人类体细胞 DNA 碱基对数目的一半。如将在第三章阐明的那样，人类是二倍体生物体，除男性的性染色体外，每条染色体有 2 个拷贝。因此，在每个人类体细胞中实际上 DNA 有 6×10^{9} bp。

2. 正如 Watson 和 Crick 所指出的那样，双螺旋互补结构直接表明 DNA 复制的机制。每一条链含有 DNA 分子的全部信息，当双螺旋解旋和复制时，每一条单链都可作为合成一条新的互补链的模板（图 2-4）。这种复制模式称为“半保留复制”，因为每个子 DNA 链包含一条亲链和一条合成的新链。

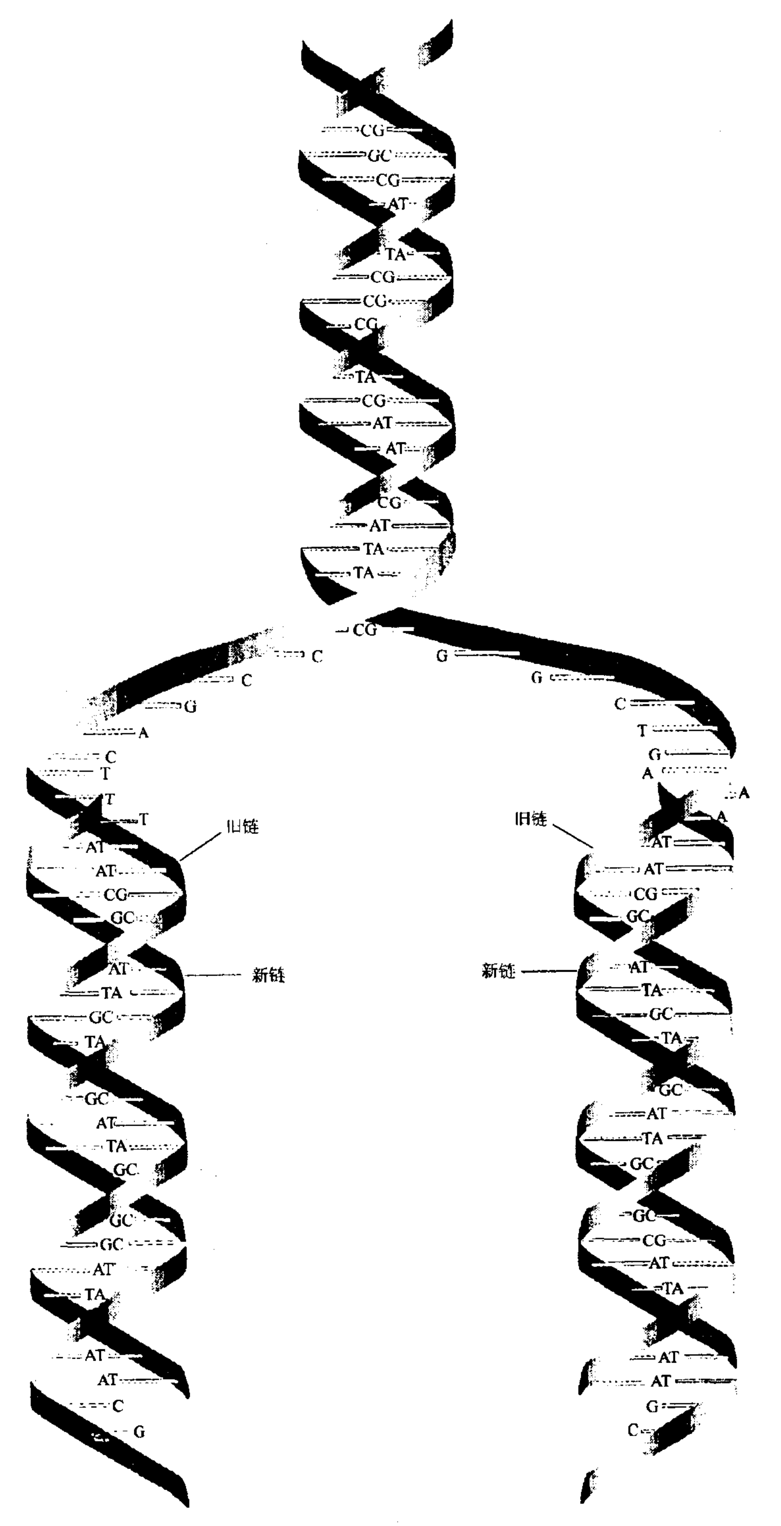


图 2-4 DNA 复制。相同的子代 DNA 分子由亲代分子解旋产生，亲代的每一条链作为模板，根据碱基配对原则合成新链。

3. 互补的结构可以为 DNA 损伤造成信息的丢失提供防护。一条链上的一个碱基受损或缺失时,它能利用互补链指导其修复而被替代。同样地,当糖 - 磷酸骨架断裂时,如果是一单链分子就几乎不能正确地重新连接,而双链分子则能修复而无相邻碱基的丢失。

4. DNA 双链的互补性使它们在复杂的分子混合物中能够彼此发现。在某些情况下,这种"再复性"或"分子杂交"的过程通过细胞核结构应用于调节基因表达。而且,正如第五章所述,这种现象已在分子生物学中广为利用,并且是其成功的本质。

三、转录成 RNA

甚至早在 19 世纪,人们就已经知道细胞中还有另一种核酸。它存在于细胞质中,特别是 与多聚核糖体(细胞的蛋白质合成工厂)有关。化学和结构分析揭示这种核酸与 DNA 相似,但有三个关键的不同(图 2-5):(a)它是单链。(b)糖是核糖而非脱氧核糖。(c)嘧啶碱基是尿嘧啶代替胸腺嘧啶。

此核糖核酸(RNA)在细胞中有很多作用。某些 RNA 称为核糖体 RNA,构成多聚核糖体的一部分,在核仁中合成。小分子的 RNA 称为转运 RNA(tRNA),是遗传信息翻译成蛋白质分子的关键元件(见下面)。其他的 RNA 与转录加工功能有关。但主要的一种令生物学家和遗传学家关注的 RNA 是信使 RNA(mRNA),它在一个基因含有的信息与其最终产物(蛋白质的特殊氨基酸序列)之间有重要的连接作用。

我们将在第五章中介绍全部转录过程,更详细地叙述分子生物学家所应用的各种技术。然而,暂时把转录考虑为在细胞核内由双链 DNA 模板产生单链 RNA 是有用的。该过程在 RNA 聚合酶Ⅱ的作用下,常由 5'→3'方向进行,即新的核糖核酸按 DNA 的指令附加到生长链的 3'端(图 2-6)。为了这一过程的成功进行,DNA 中必须存在一些信号(启动子)以指明转录从哪里开始。同样,转录并不是一直进行下去,而必须在所需序列被转录后停止。

DNA

鸟嘌呤

胸腺嘧啶

RNA

鸟嘌呤

尿嘧啶

图 2-5 DNA 和 RNA 化学结构的比较。DNA 序列为 5'-GT-3',RNA 序列为 5'-GU-3'。重点是 RNA 核糖环中额外的 2'羟基。DNA 中的碱基 T 在 RNA 中被 U 代替,U 缺少嘧啶环上的甲基。并未显示额外的区别——通常 DNA 为双链而 RNA 为单链。

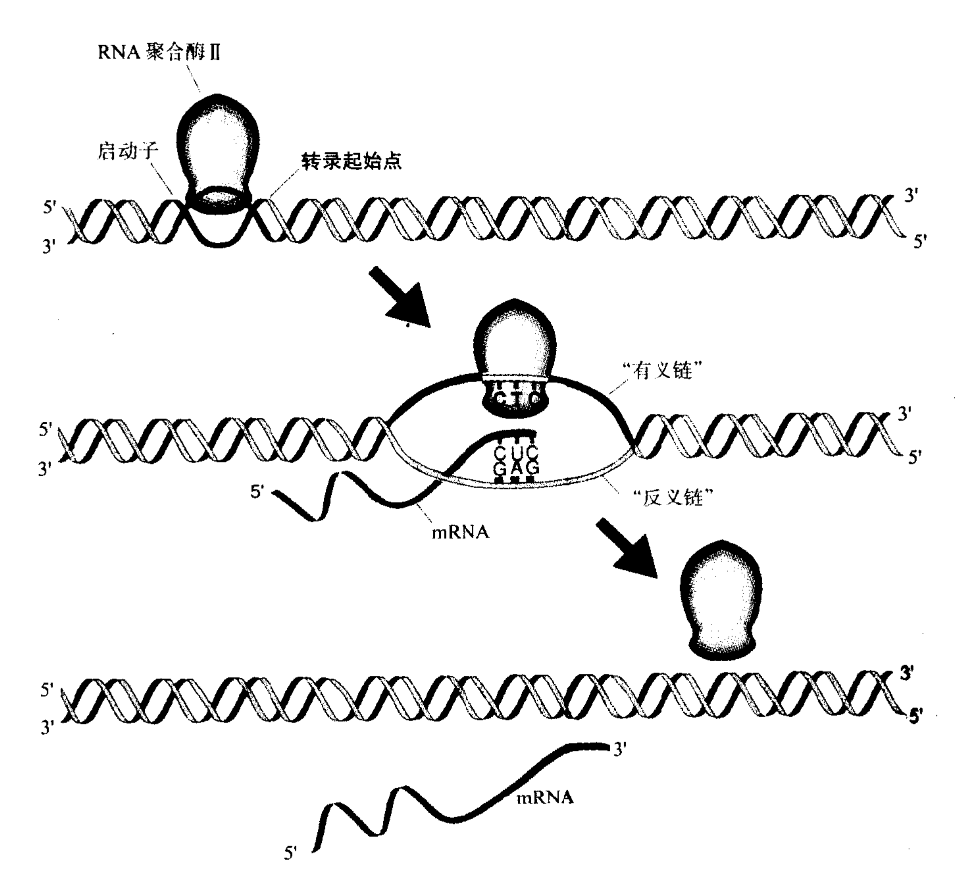


图 2-6　转录过程的示意图。RNA 聚合酶Ⅱ识别基因 5′端的特异序列(启动子),并开始将它转录成信使 RNA(mRNA)。mRNA 从 5′→3′方向合成,与 5′→3′的 DNA(又称"有义链")有相同序列。RNA 的形成机制依赖于作为复制模板的 DNA 反义链与新形成的 RNA 间的碱基配对。

四、翻译成蛋白质

由转录生成的 mRNA 必须转移到核外的核糖体,在核糖体被翻译成蛋白质序列。近 30 年来的最大成果之一是阐明了翻译的分子机制及用于这一过程中的"语言"(通常称为遗传密码)。因为在 DNA 和 RNA 中只有四种碱基,而蛋白质却有 20 种氨基酸,所以至少应该需要 3 个碱基($4^2=16,4^3=64$)明确地指定一个氨基酸。在 1961 年到 1966 年之间进行的一系列精细的试验推测出正确的三碱基单词或"密码子",并导致遗传密码的确定,如表 2-2 所示。值得注意的是,在此遗传密码中存在"兼并性",所以数个密码子可产生相同的氨基酸。例如密码子 UUA、UUG、CUU、CUC、CUA 和 CUG 都编码亮氨酸。三个密码子,UAA、UAG 和 UGA,不编码氨基酸,但导致翻译停止,称为终止密码子。

在核糖体上执行的翻译过程是复杂的。一批氨基酸特异的受体分子(称为转运 RNA)共价地结合一个特定的氨基酸。此三碱基"反密码子"与特定的氨基酸密码子互补,并使之精确地加到生长的多肽链上,如图 2-7 所示。这个过程一直延续至到达终止密码

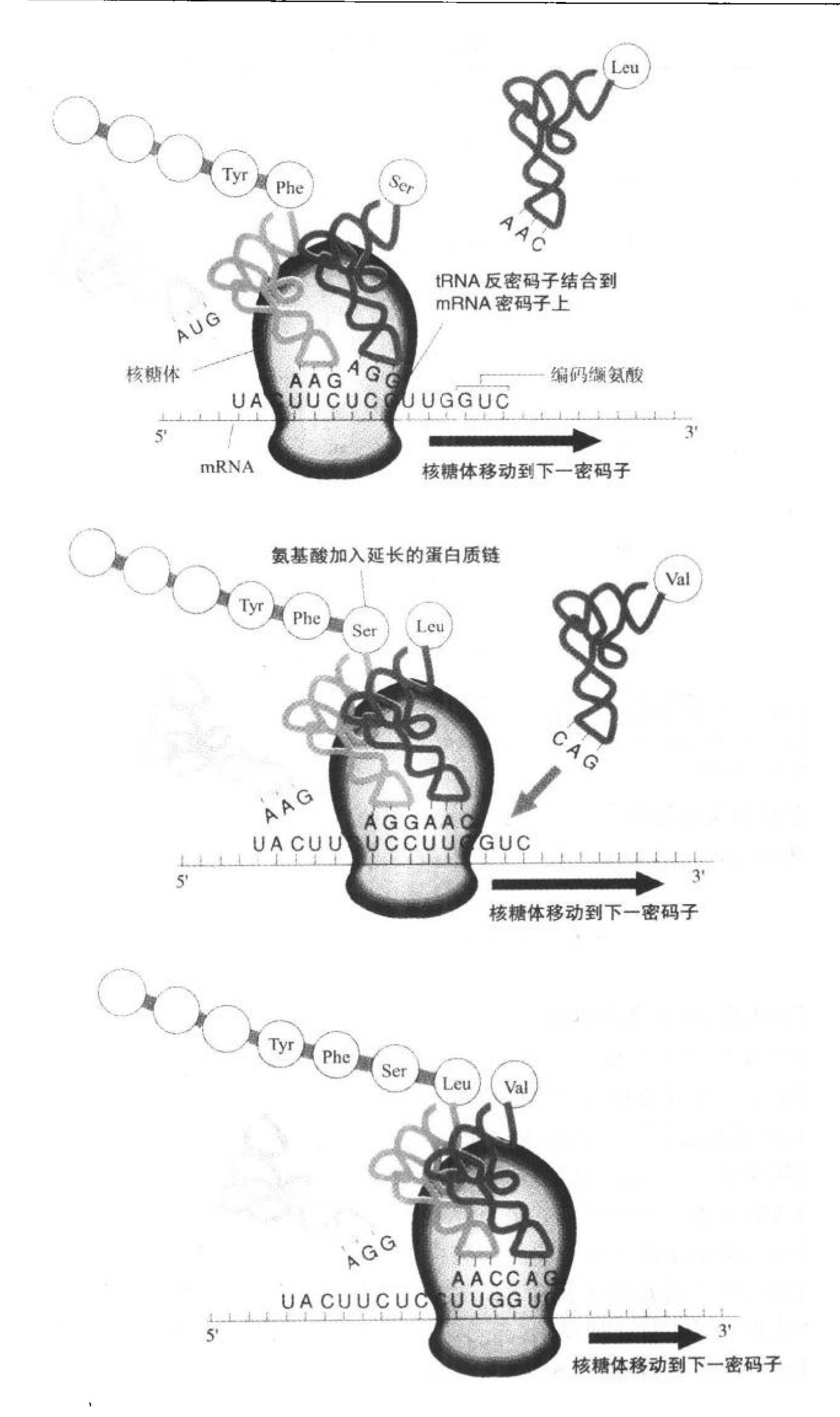


图 2-7　通过核糖体将 mRNA 翻译成蛋白质。以 tRNA 及它们的反密码子作为受体，mRNA 的 5′-UACUUCUCCUUGGUC-3′顺序被翻译成酪氨酸－苯丙氨酸－丝氨酸－亮氨酸－缬氨酸的氨基酸顺序。

表 2-2　遗传密码[a]

第一位置(5'端)	第二位置 U	C	A	G	第三位置(3'端)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	终止密码	终止密码	A
	Leu	Ser	终止密码	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

a. 氨基酸缩写为：Ala 丙氨酸；Arg 精氨酸；Asn 天冬酰胺；Asp 天冬氨酸；Cys 半胱氨酸；Gln 谷氨酰胺；Glu 谷氨酸；Gly 甘氨酸；His 组氨酸；Ile 异亮氨酸；Lys 赖氨酸；Met 蛋氨酸；Phe 苯丙氨酸；Pro 脯氨酸；Ser 色氨酸；Thr 苏氨酸；Trp 色氨酸；Tyr 酪氨酸；Val 缬氨酸。

子，导致多肽链从核糖体上释放。其他氨基酸的序列是 mRNA 的 5'端对应于蛋白质的氨基(NH_2)末端，而 mRNA 的 3'端对应于蛋白质的羧基(COOH)末端。

五、DNA 的高度螺旋次序

很可能人类 46 条染色体的每一条都是由双链 DNA 的单个分子组成。若伸展开来，一个细胞中的 DNA 可长达大约 2m。显然，细胞必定应用某种有效的包装方法来处理如此复杂的分子。哺乳动物细胞中存在一个精确的螺旋化系统，该系统可能也参与基因表达的调控，但仍知之甚少。以称为组蛋白的基础蛋白为核心，由大约 146bp DNA 组成的双环缠绕其周围(图 2-8)。这个单位称为核小体。组蛋白结构在进化中的高度保守性具有重要的功能作用。这种“串珠”状 DNA 结构导致长度压缩大约 7 倍，进一步通过以螺线管方式排列的核小体及更高级的复杂结构进行组织(图 2-8)。

基因表达调控相关的因子与该单一组蛋白 - DNA 复合体相互作用的机制是当前研究的一个重要领域。有证据显示，参与调控的 DNA 区域可能局部没有组蛋白，例如，活性表达基因的一定区域(特别是 5'端)通常对切割 DNA 的酶(核酸酶)的切割非常敏感，但是该酶并不能有效地切割组蛋白结合的 DNA。

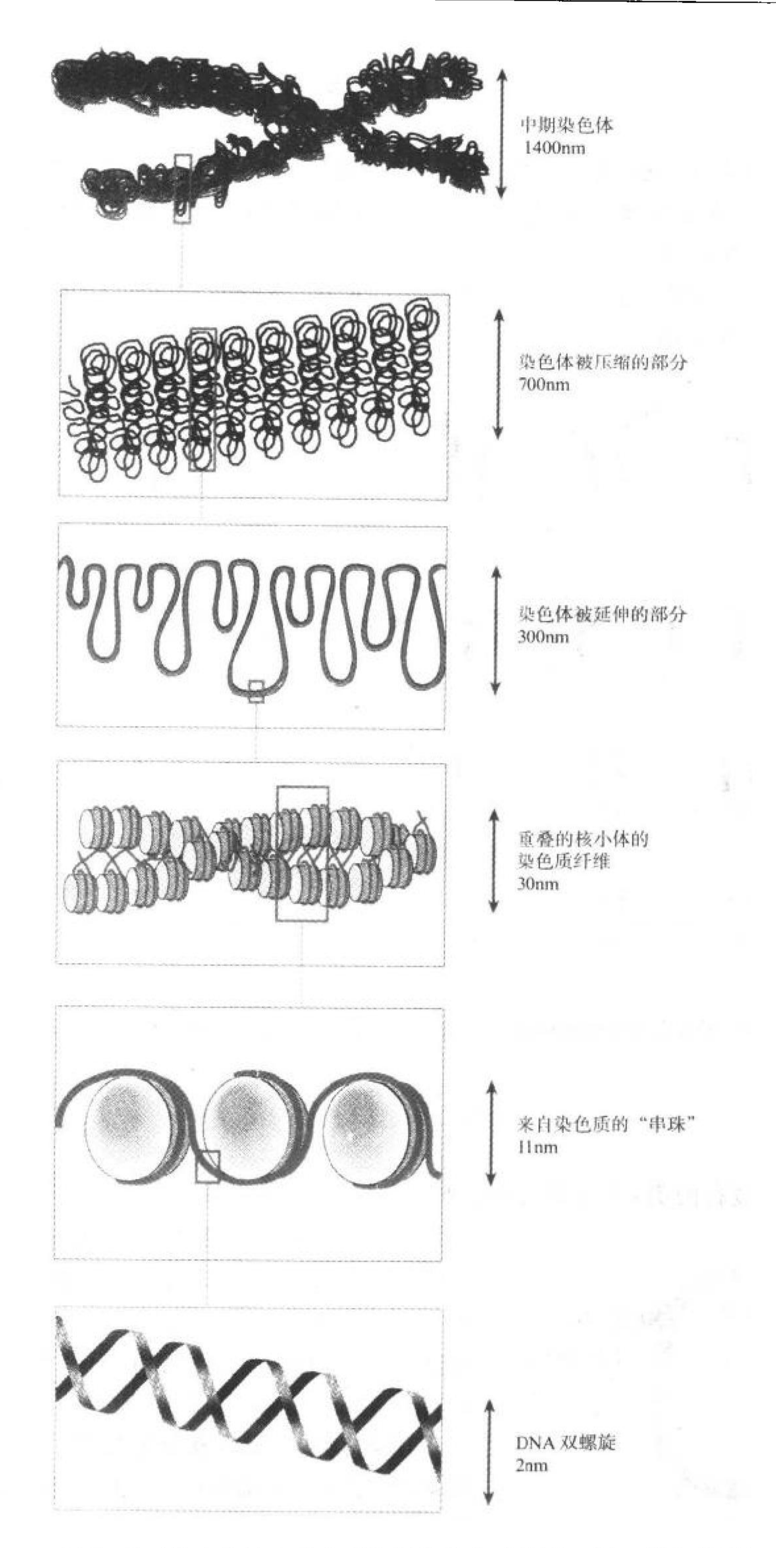


图 2-8　DNA 螺旋,从下往上按越来越高度组织化的方式排列。每一部分都以(nm)的形式显示大小。

六、人类染色体

最高度螺旋的DNA为染色体。每一物种有特定数目和大小的染色体，称为核型。人类的核型由46条染色体组成，这将在第八章更为详细地讨论(图2-9)。如上所述，人类是一个二倍体的生物体，其中44条(常染色体)由22对同源的常染色体组成，每对中的每一条都含有相同的遗传信息。其余两条是性染色体，女性有两条X染色体，而男性有一条X染色体和一条Y染色体。

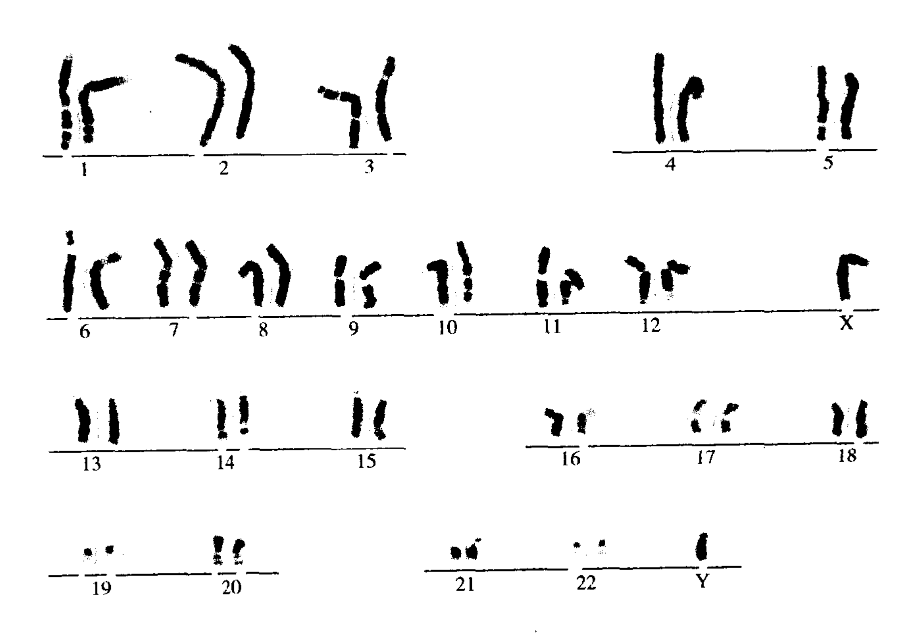


图2-9 正常人类男性的核型。由22对常染色体，一条X染色体和一条Y染色体组成。

七、有丝分裂

细胞分裂有两类：有丝分裂和减数分裂。有丝分裂是一个细胞分裂产生两个在遗传上与亲代完全相同的细胞的过程。有丝分裂使一个受精卵发育成为一个约有10^{14}个细胞的完全的人体，所有这些细胞在遗传上与原始单个细胞完全相同(少数例外)。在有丝分裂中，每个子代细胞一定接受全部的46条染色体的成分。

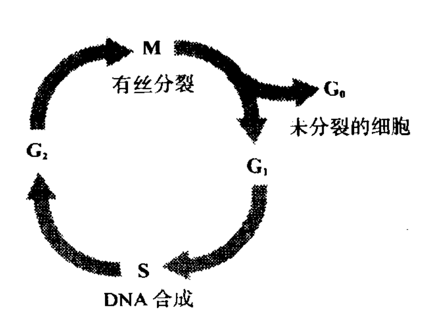


图2-10 细胞周期。

有丝分裂本身即核分裂过程，只需很短的时间。然而，它却是一个精细程序化的过程，如图2-10所示，称为细胞周期。分裂刚刚结束后，注定再分裂的细胞进入所谓的G_1期，而不分裂的细胞进入所谓的静止期即G_0期。G_1期的细胞下一步进入S期。在此时

期,如前所述,DNA 通过半保留的机制发生复制,每条 DNA 单链作为它自身复制的模板。

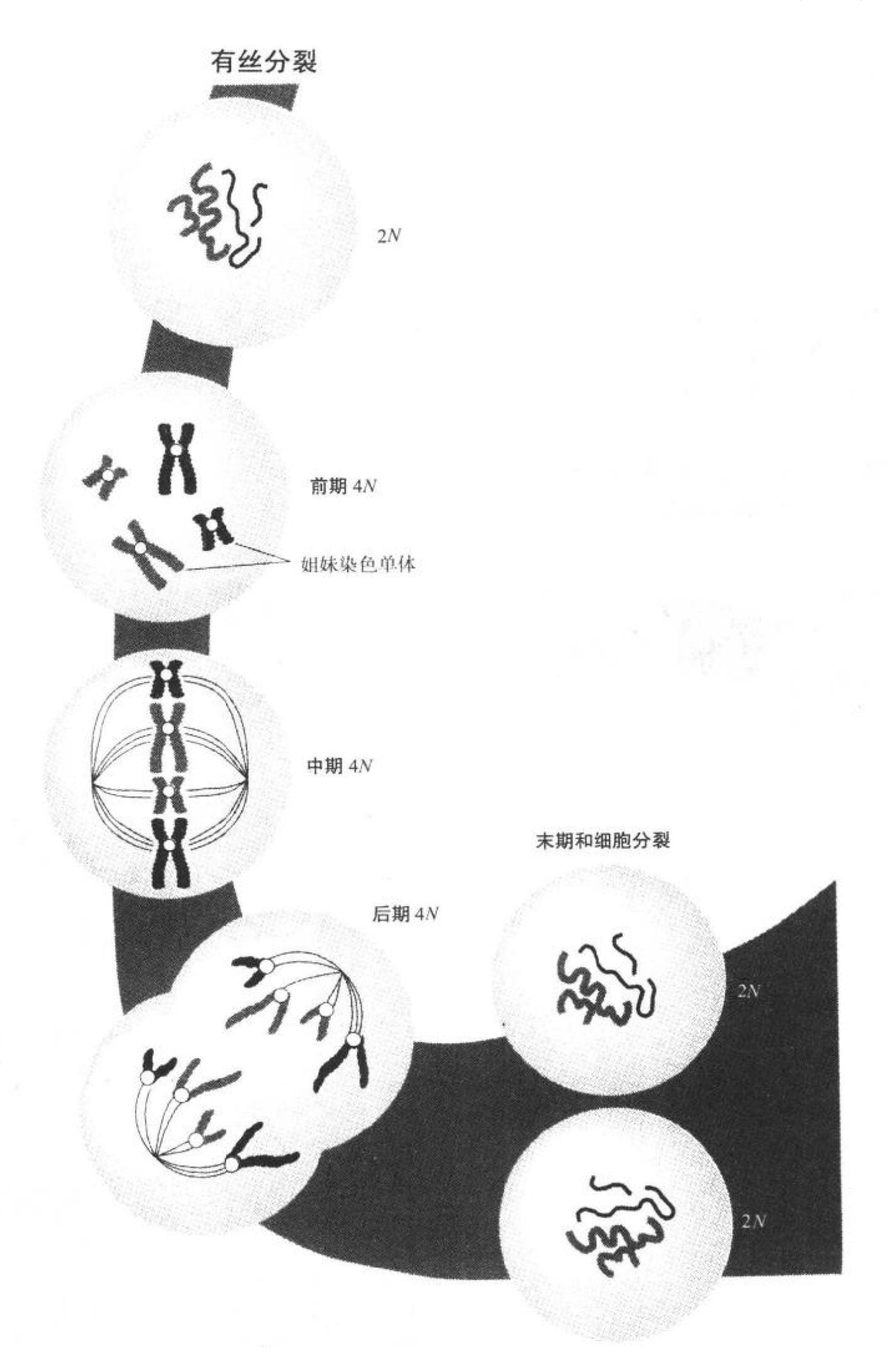


图 2-11　有丝分裂。为了简便,仅以由 2 对常染色体组成的 4 条染色体为例。每一期都有细胞的"倍体性"表示,2N 代表二倍体状态。在形成姐妹染色体单体之后但在细胞分裂之前,细胞含有 4N 的 DNA 量。

结果46条染色体的每一条都复制成"姐妹染色单体",由称为着丝粒的中心缩窄部分相连。S期结束,另一间隙期(G_2期)开始,它引导进入真正的有丝分裂(M)。

在有丝分裂(图2-11)中,姐妹染色单体和着丝粒清晰可见,并沿最终要分裂的赤道板排列。然后,所有46条染色体的着丝粒分裂,以致每一条姐妹染色单体都以进入每个子细胞中而告终,完成细胞周期。

八、减数分裂和配子发生

减数分裂是另一类细胞分裂,用来产生男性和女性的配子(分别是精子和卵子)。它与有丝分裂在决定性方面不同。极少数反映认为必须如此,即如果精子和卵子含有全套完全互补的("二倍体")46条染色体,那么受精卵将含有96条染色体,包括3条X和一条Y染色体!

减数分裂是一特殊的减数过程,即在配子发生过程中产生精子和卵细胞,每个都只有23条染色体("单倍体"状态),如图2-12所示。这实际上有两次分裂,减数分裂Ⅰ期和

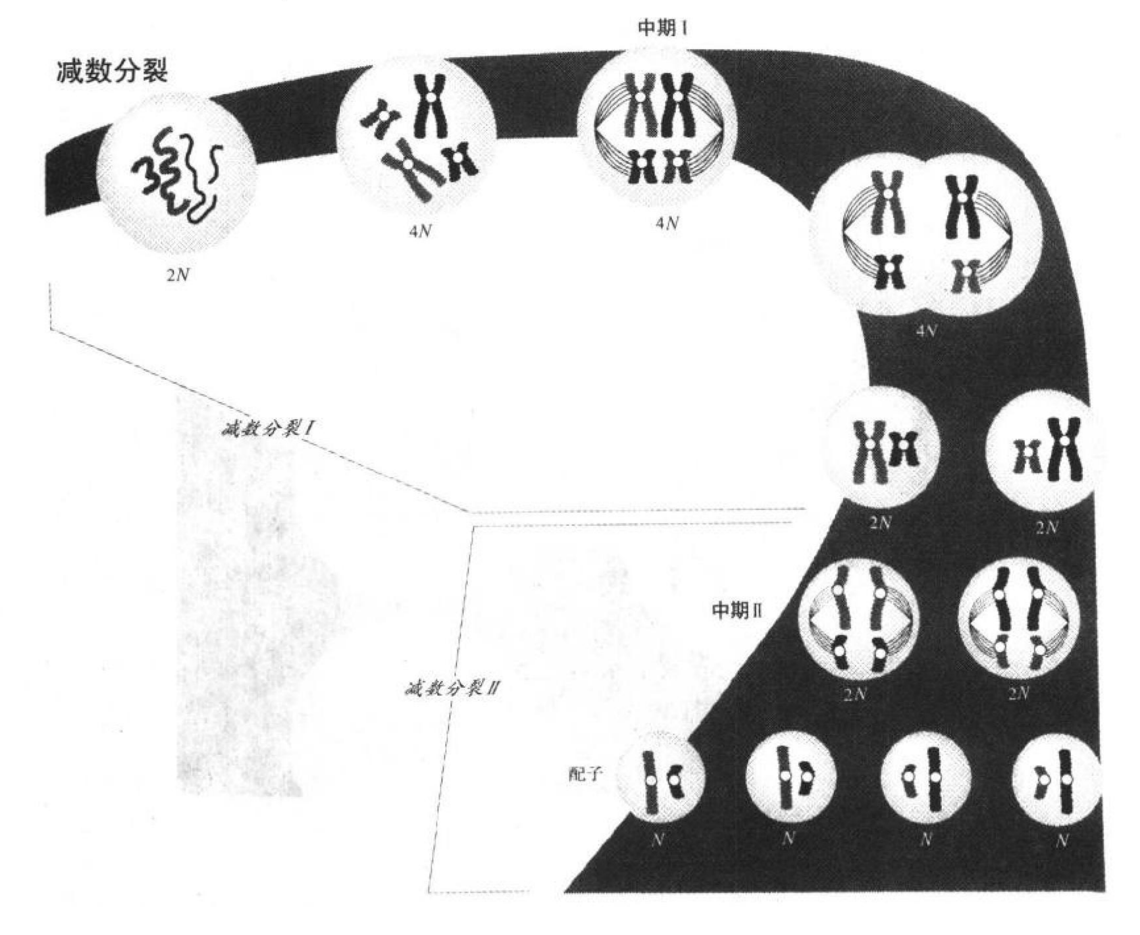


图2-12 减数分裂。也只用两对常染色体表示。注意,减数分裂Ⅰ期后,每个细胞只剩一个同源染色体。然后,减数分裂Ⅱ期导致姐妹染色体单体分开。为简单描述,没有表示交换。

Ⅱ期。在减数分裂Ⅰ期,如同有丝分裂一样,每条染色体复制成姐妹染色单体。但与有丝分裂不同的是,同源染色体成对排列(这一过程称为联会),并分离到相对的两极,姐妹染色单体仍在一起。在减数分裂Ⅱ期,姐妹染色单体分离,导致每个配子中含有23条染色体。

九、交　换

减数分裂Ⅰ期的一个非常重要的特征是在联会过程中,即当同源染色体配对时,发生交换(图2-13)。特殊的结果是在减数分裂结束时,在一个配子中保留的染色体可能是有两个同源亲代染色体的拼凑物。然而,一条染色体上紧密相连的基因(如图2-13的A和B)可能一起传递,而在染色体上距离远的基因之间可能发生交换(或"重组")。在减数分裂过程中平均大约发生30~40次交换(或每1~2次/染色体)。

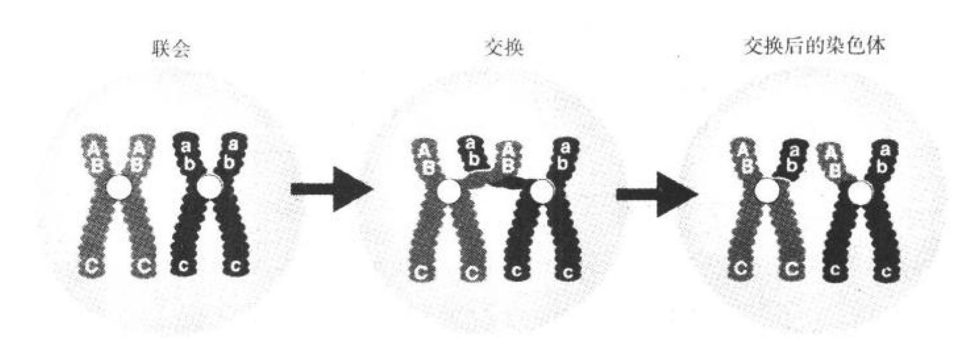


图2-13　减数分裂Ⅰ期的交换。

交换除极大地增加配子潜在的遗传多样性之外,还可提供在同一条染色体上两个分离的基因(称为"同线的"基因)间距离的数量评估。正如我们将在第九章中看到的,交换为定位基因提供了一个有力的手段。

十、突　变

直到现在我们仍将组成人类基因组的DNA看成是不变的和没有差错的。然而,若是这样的话,进化将是不可能的。换言之,DNA序列发生突变,一定会使生物体逐渐发生选择性改变,但也不会太频繁,因为大部分突变将是有害的。正如指出的,以未受损的DNA链做模板改正受损的DNA可纠正很多突变。但是,若损伤发生在两条链或发生在DNA复制之前或复制过程中,造成的DNA序列的变化对该细胞所有未来的后代是永久性的。

人们已尽力去试图确定人类的背景突变率(background mutation rate)。当基因从亲代传给子代时,一个给定基因的新突变频率约为百万分之一。因为估计人类约有80 000~100 000个基因,所以,这表明所有新生儿的8%~10%带有一个新突变。幸运的是其中大多数都是静止的,因为同源染色体上等位基因是正常的,因此突变的存在并不表现。

在分子水平上，新突变可有各种形式。正如我们将在这本书中所看到的，在DNA水平上，增加的突变数目的确定已总结出规律："无论什么错误，终究总会发生。"（基因组的Murphy法则?）。图2-14列举了几个例子。最简单的变化是一个核苷酸替代另一个，称为"点突变"。这若发生在编码区并改变三联体密码子，使其成为编码一个不同的氨基酸的密码子（"错义突变"），则产生的蛋白质将有结构上的不同。如果氨基酸替代发生在重要功能区，或改变蛋白质的结构或稳定性，这必将导致功能的改变。一个特殊的破坏性点突变是正常的密码子变成终止密码子UAA、UGA或UAG，产生一个提前终止密码子，这种突变称为"无义突变"。

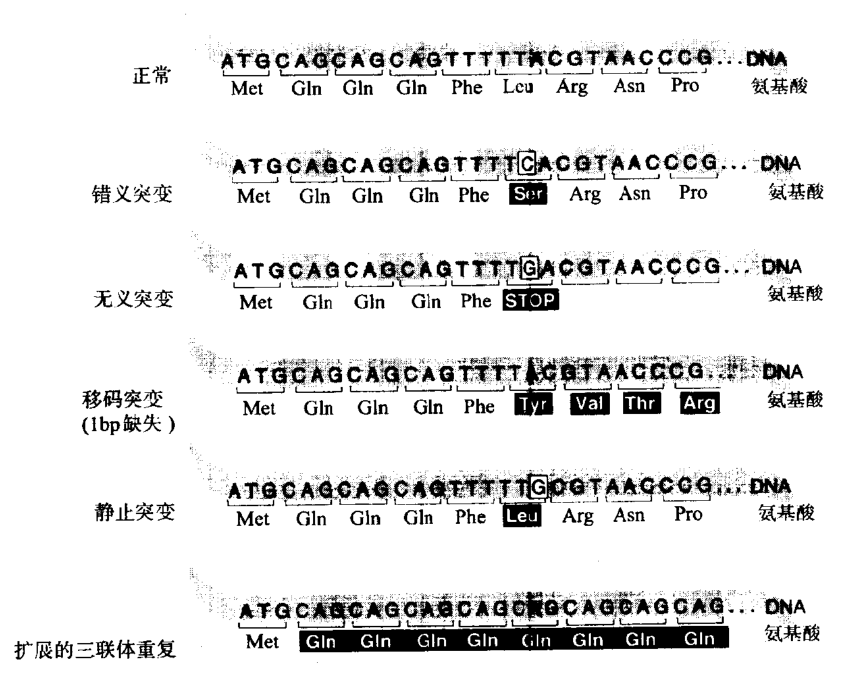


图2-14　突变的例子。图示"有义链"DNA的编码序列和编码的氨基酸序列。4种不同的突变影响一个亮氨酸的密码子，同时举例表示一个扩展的正常的$(CAG)_3$序列。

若编码区内插入或缺失非3的倍数的少数核苷酸，则从该点向后翻译的可读框架将会改变，导致一个完全异常的蛋白质羧基末端，这种突变称为"移码突变"。

还有其他种类的突变影响一个基因转录，我们将在第五章讨论，那时将更为全面地描述基因的解剖学。首先，我们将深入探讨产生人类疾病的突变效应及其所产生的遗传方式。

推荐读物

分子生物学史

Judson HF. The eight day of creation: makers of the revolution in biology. New York: Simon & Schusrer, 1979.

Watson JD. The double helix. New York: Athenum, 1968.

普通参考教材

Alberts B, Bray D, Raff M, Roberts K, Watson JD. Molecular biology of the cell. 3rd ed. New York: Garland, 1994.

Darnel J, Lordish H, Baltimore D. Molecular cell biology. 2nd ed. New York: Scientific American Books, WH Freeman, 1990.

Watson JD, Tooze J, Kurtz DT. Recombinant DNA, a short course. New York: Scientific American Books, WH Freeman, 1983.

Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM. Molecular biology of the gene. 4th ed. Menlo Park, CA: Benjamin/Cummings Publishing Co., 1987.

经典论文

Meselson M, Stahl FW. The replication of DNA in *E. coli*. Proc Acad Sci USA 1958;44:671 ~ 682.

Nirenberg MW, Matthaei JH. The dependence of cell-free protein synthesis in *E. coli*. upon naturally occurring or synthetic polyribnucleotides. Proc Acad Sci USA 1961;47: 1588 ~ 1602.

Watson JD, Crick FHC. A structure for deoxyribose nucleic acid. Natrue 1953;171:737 ~ 738.

第三章　孟德尔式遗传

那些能够完全遗传的或者通过杂交几乎不变的，并因而构成杂种性状的特性称为显性，而在此过程中成为潜在的那些特性则称为隐性。

——Gregor Mendel，1865

为了理解孟德尔式遗传，首先要明确几个基本术语。**遗传基因座**（genetic locus）是指位于一条染色体上的特定位置。基因座通常用来指一特异的基因。**等位基因**（alleles）是在一个既定位点上一个基因或 DNA 序列的替换形式。如果在一个基因座上的两个等位基因是相同的，则此个体在该基因座是**纯合性的**（homozygous）；如果它们是不同的，他或她是**杂合性的**（heterozygous）。这样的个体分别称为**纯合子**（homozygote）和**杂合子**（heterozygote）。在一个既定基因座上有两个不同突变等位基因的个体称为**复合杂合子**（compound heterozygote），而在两个不同基因座上各有一突变等位基因的个体称为**双杂合子**（double heterozygote）。

基因型（genotype）是一个个体的遗传结构或组成，也可指一个特定的遗传基因座上的等位基因。**表型**（phenotype）是基因型和环境因素相互作用所观察到的结果，尤其是一个或一些特异基因的可见的表现。基因型和表型的意义可用音乐比拟来表明。图 3-1 是莫扎特协奏曲的核心部分，钢琴和管弦乐 A 大调 K.488 乐章。这些乐谱包含了每种乐曲演奏的音符和各个部分在时间上的相互关系所必须的全部信息和指令。这是基因型。表型是我们听到的声音，而声音受到环境有力的影响，包括独唱者、指挥者、管弦乐以及演出大厅或录音效果等。

我们确定孟德尔式遗传病为**单个突变的基因**引起的疾病，突变基因极大地影响表型，并以与孟德尔所描述豌豆确定分离的特征相似或相同的简单方式遗传。如果编码疾病的基因位于 22 对常染色体或非性染色体的其中一条之上，则这类孟德尔式疾病称为**常染色体病**；如果由位于 X 染色体上的一个突变基因编码，则称为 **X 连锁遗传病**。按孟德尔定律，我们将在杂合子表达的那些疾病确定为**显性**（dominant），杂合子是指有一个突变等位基因拷贝和一个正常或野生型等位基因拷贝的个体。而**隐性**（recessive）为只是突变等位基因的纯合性个体（或是两个不同的突变等位基因的复合杂合子），即携带双倍量的异常基因的个体才在临床上表现症状的那些疾病。应强调的是显性和隐性是指性状或表型而不是基因。虽然我们有时说显性和隐性基因，但这是一种速记，而应理解为指的是性状。到目前为止，已编目的并已知按孟德尔式遗传的人类表型超过 5 000 种。一半以上是常染色体显性性状，36% 是常染色体隐性性状，X 连锁遗传不足 10%。在这 5 000 种性状中，大约 4 000 种与人类疾病有关，其中几乎 600 种已鉴定有一种或几种致
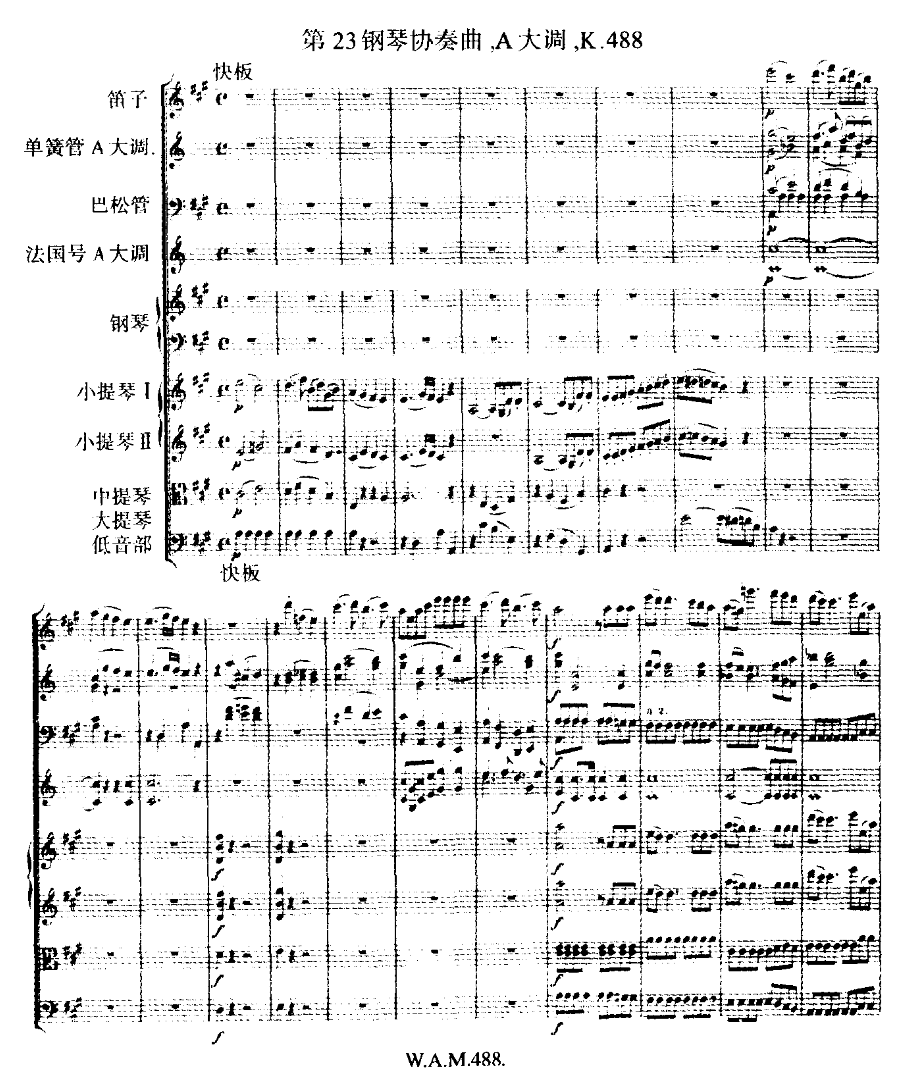


图3-1 莫扎特协奏曲的管弦乐核心部分第一页,钢琴和管弦乐A大调K.488乐章(引自Mozart WA. Piano Concertos Nos. 23 ~ 27 in Full Score. New York: Dover Publications, 1978:1.)。

病突变。

大多数孟德尔式性状的遗传方式是通过观察这些性状在家系内的分离或传递而推断的,这些信息通过系谱形式表达。系谱一词来源于法语“pied ed grue”或鹤足(crane's foot),以示图解的分支模式。用于绘制人类系谱的标准符号见图3-2。考虑到近年来医学遗传学中的技术进步,已公布了更广泛推荐标准的人类系谱命名法。先证者或索引病例(index case)是第一个前来就诊的受累家族成员,用箭头和字母P表示;仅有箭头表示咨询者或寻求遗传咨询或测试的个体。系谱中个体所处的世代位置以罗马数字表示,阿拉伯数字则表示单一世代中的个体。在医生的记录或医院的图表的系谱内,个体的年龄记录于符号的旁边或下面而不是里边。

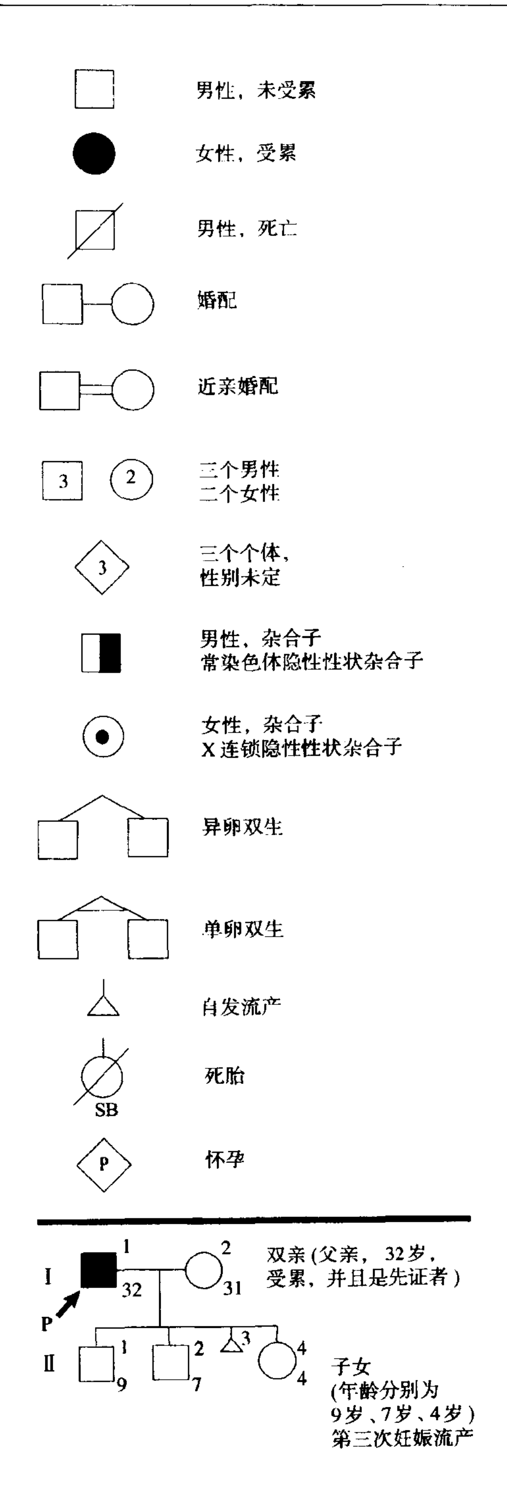

男性，未受累
女性，受累
男性，死亡
婚配
近亲婚配
3
2
三个男性
二个女性
3
三个个体，
性别未定
男性，杂合子
常染色体隐性性状杂合子
女性，杂合子
X连锁隐性性状杂合子
异卵双生
单卵双生
自发流产
死胎
SB
怀孕
P
I
1
32
2
31
双亲(父亲，32岁，
受累，并且是先证者)
P
II
1
9
2
7
3
4
4
子女
(年龄分别为
9岁、7岁、4岁)
第三次妊娠流产

图 3-2 标准系谱符号。

一、常染色体显性遗传

常染色体显性遗传病包括许多严重的和较常见的成年遗传病，诸如家族性高胆固醇血症、遗传性结肠癌、多囊肾病、Huntington病和神经纤维瘤。一个家族性高胆固醇血症家系的系谱如图3-3所示。这种疾病的特征是血中胆固醇水平增高，以及早熟型动脉粥样硬化性心血管疾病（在生命早期发生心肌梗死或突发心脏病）。虽然男女均可受累，但男性患冠状动脉疾病的倾向比女性约早10年，在无高胆固醇血症的一般群体中也是这样。已经证实此病的基本缺陷是细胞表面低密度脂蛋白受体的缺乏（见第七章）。W.H.（图3-3系谱中Ⅱ-4）是一位33岁的白人广告商。25岁时开始反复发作胸痛，30岁时患心肌梗死。冠状动脉造影显示存在广泛的冠状动脉疾病，他接受了冠状动脉旁路移植手术。身体检查显示胆固醇在伸肌腱沉积的迹象，实验室检查提示严重的高胆固醇血症。家族史表明他的父亲52岁时死于第二次心肌梗死，他的伯父（Ⅰ-1）46岁时也死于心肌梗死，他50岁的姐姐无症状，但患有高胆固醇血症。W.H.的三个孩子完全健康，身体检查正常，但都有严重的高胆固醇血症。

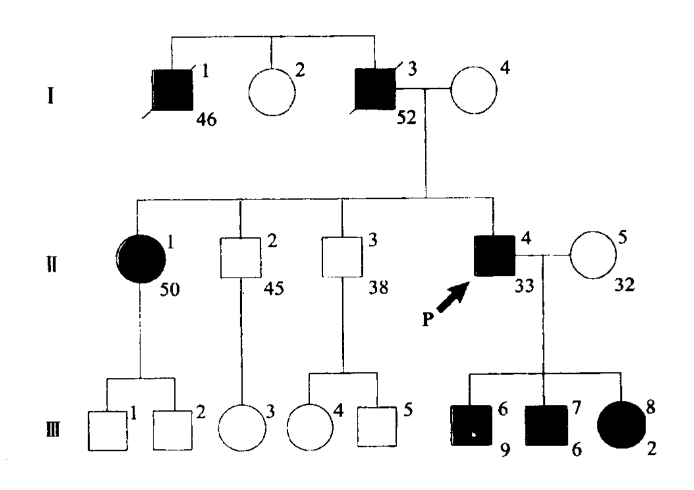
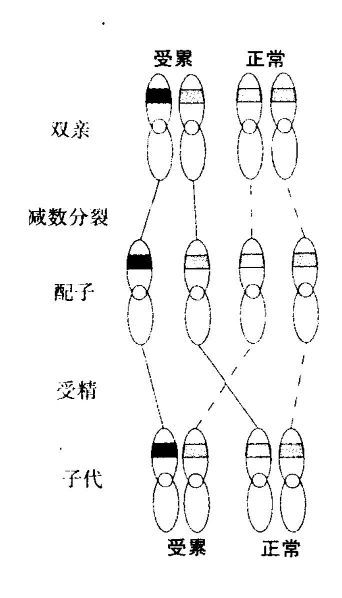


图3-3 常染色体显性家族性高胆固醇血症家系的系谱受累个体用实心黑色标记表示。

正如此系谱所见，遗传方式的特征是垂直的，即疾病是以垂直方式从一代传递到下一代。虽然表现的严重程度可受个体性别的影响，但男性和女性均受累，并以相同的概率传递其性状。受累个体的双亲之一是患者。注意：在显性遗传疾病的系谱中，未受累的配偶常被省略。

常染色体显性性状的遗传方式通过观察减数分裂过程中携带突变等位基因和正常等位基因的染色体分离就一目了然（图3-4）。为了简明易懂，假定只有一个正常等位基因和一个突变等位基因，可能有三种基因型：正常纯合性，受累纯合性和杂合性。因为男女均可能具有这三种基因型的任何一种，所以可形成6种婚配型。然而，由于男女可以相同的概率传递异常基因，所以性别差异可忽略。因为纯合受累个体很罕见，所以在常



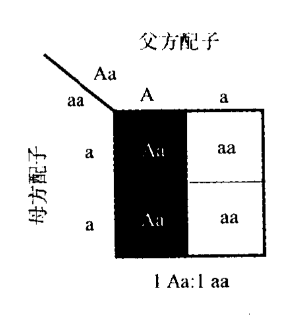


图 3-4 常染色体显性遗传。婚配图解(上)和 Punnett 方格图(下)表明杂合受累个体和纯合正常个体的婚配结果。在婚配图解中突变的等位基因以黑色标明，在 Punnett 方格中受累的后代以黑色标明。

染色体显性遗传病中通常是在正常纯合子个体和杂合受累个体之间的婚配型。由这些个体产生的配子将仅含此对基因的一个等位基因，故可能产生的配子如图 3-4 上部婚配图解所示，图 3-4 下部的 Punnett 方格图是用另一种方式表示产生的配子和受精时他们可能的结合。这样婚配的每一个子代有 50% 的概率遗传携带突变等位基因的染色体，50% 概率遗传携带正常等位基因的染色体。因此，有 50% 的受累概率。受累个体的子代平均约有 1/2 发病。未受累的孩子获得携带正常等位基因的染色体，只会有正常的后代。例如，在图 3-3 所示的系谱中，受累亲代Ⅰ-3 的后代 4 个孩子中有 2 个受累，即可预期 50% 的概率。然而，在任何特定的一群同胞(一组同胞)中，可能不符合 50% 的预期值，如 W.H 的 3 个后代(Ⅲ-6、Ⅲ-7、Ⅲ-8)均患有家族性高胆固醇血症。

为理解这种与预期结果的偏差，回忆每一子代代表受累概率为 50% 的独立事件是重要的。按照**独立原则**(independence principle)，两个或更多的独立事件联合发生的概率是它们独立概率的乘积。因此，所有三个孩子都受累的概率是其独立概率的乘积，即 $(1/2)\times(1/2)\times(1/2)=1/8$。

总之，常染色体显性性状或疾病的典型系谱模式是十分明显的，表现为垂直遗传，男女发病机会均等，双亲之一受累的子代平均 1/2 发病，未受累双亲的后代则无人发病。事实上，一种罕见的疾病或性状的这种遗传方式通过三代传递即可确定该性状是按常染色体显性遗传方式遗传。

(一) 变异问题

在常染色体显性遗传中，除全部受累个体都有双亲之一受累的规律之外，还有一些特殊的表现。在这些情况下，常染色体显性遗传病患者可能是散发病例，也就是说，他们的双亲显然都是未受累的。首先，散发病例可能是由一个新的**突变**或遗传物质的改变所引起的。一种与生育力有关的疾病越严重，该病的患者就越可能是由形成这种疾病的配子之一发生新的突变所致。如果常染色体显性疾病通常引起在生育年龄前死亡或生育障碍，那么这样的病例必定是由于一个新突变所致。Ⅱ型成骨不全(在第七章讨论)是一种产期致死性常染色体显性疾病，所有的病例都是新的突变。另一个例子为软骨发育不全，它是短肢性侏儒症的一种 常见类型(图 3-5)，生育适合度明显地降低，80% 以上的病例属新的自发突变。有证据表明，父亲年龄的增高可能增加突变风险。软骨发育不全中突变的基因编码Ⅲ型成纤维细胞生长因子受体，这种受体介导成纤维细胞生长因子对软

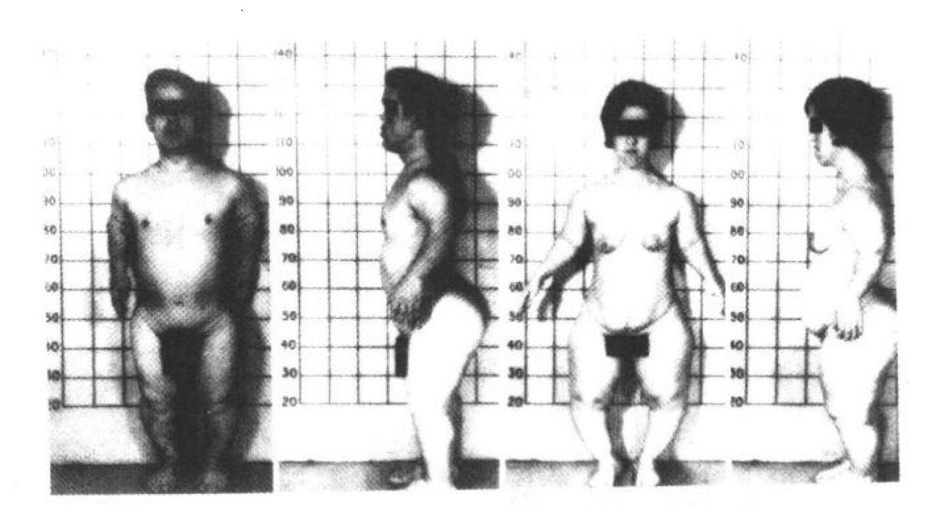


图 3-5 软骨发育不全的男人和女人。体矮,头大,四肢短,尤其是近端部分,脊柱弓形前凸(引自 Mckusick VA. Hevitable disorders of connective tissue. 4th ed. St Louis: CV Mosby, 1972:758.)。

骨的效应。奇怪的是,几乎所有受累的个体都有这个基因的完全相同的核苷酸突变,这在人类遗传病中实属罕见。更为常见的是多个独立的突变引起一种疾病(这一点将在本章后面和其他章节中讨论)。

第二种情况是患儿的双亲表现正常,它可能由于**外显率降低**(decreased penetrance)所致。外显率表现为全或无的现象,是指基因突变在临床上的表达或缺乏。在某种意义上讲,外显率是我们认识突变基因的表达或表型的能力的一种人为现象。当确定表达是临床疾病时,携带突变基因的个体可能不表现疾病,那么该病是非外显的。我们从量上来确定外显率是通过确定肯定基因携带者(杂合子)的比例来实现的,因为一个突变等位基因表达其表型。图 3-6 是遗传性红斑性肢痛病家系的系谱,这种罕见疾病的特征是红、热、足痛和偶尔手痛。从系谱中可清楚地看出,该病是常染色体显性遗传。系谱中Ⅰ-4、Ⅱ-3、Ⅱ-5、Ⅲ-1、Ⅲ-3、Ⅲ-7 和Ⅳ-1 ~ Ⅳ-4 等个体是突变等位基因的杂合子,除Ⅲ-1 外全部受累。这位妇女(Ⅲ-1)肯定是杂合子,因为她有一个受累的母亲和 3 个受累的孩子,但是她本人既无发病史也无该病的任何症状。在这个系谱中,10 个肯定杂合子个体中 9 人发病,表明外显率约为 90%。如果一种疾病的外显率达 90%,这意味着携带突变等位基因的个体表达该病的概率为 90%。然而这种表达本身是全和无的关系。

第三,若干重要的常染色体显性疾病在出生时并不表现,而仅在生命的后期才表现出来(延迟发病)。多囊肾病的特征是双侧肾肿大并有许多囊肿,血尿,高血压,腹痛和进行性肾衰竭。许多病人通常直到 40 岁或更晚才发生症状。直到 20 岁肾脏这种特征性囊肿才能用超声和放射学的技术检测出来。更为明显的是 Huntington 病,一种进行性神经性疾病,其特征为运动异常(舞蹈病),进行性精神功能丧失(痴呆)。如图 3-7 所示,该病平均发病年龄大约在 35 岁。因而,有可能携带致病基因的个体在该病出现症状之前生育孩子并且死于无关的原因(例如汽车肇事)。因此显示该病是隔代遗传,即受累的子代

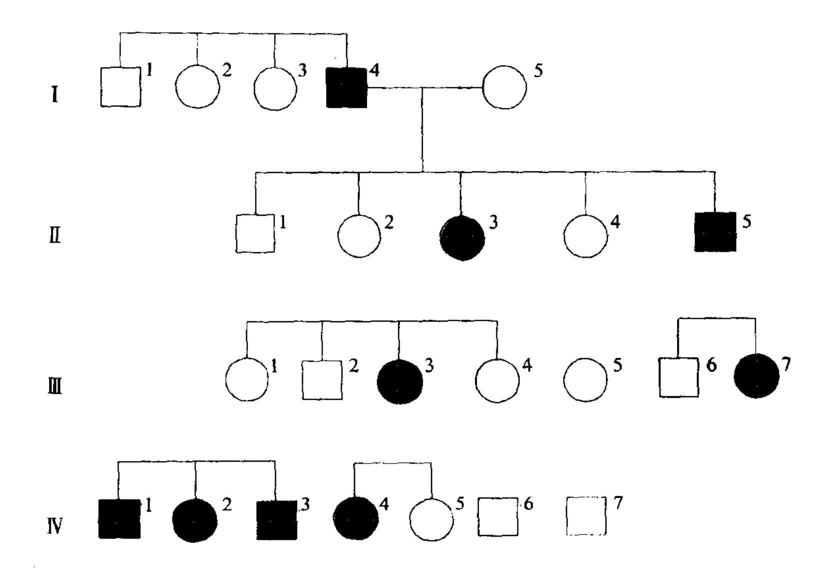


图 3-6 常染色体显性红斑性肢痛病家系系谱，表明不外显。系谱中Ⅲ-1 临床表现正常，但是一个突变等位基因的肯定杂合子。

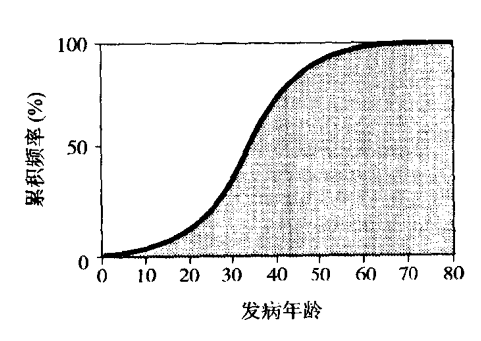


图 3-7 Huntington 病发病年龄。在这项研究中 50%的患者在 35 岁发病(引自 Conneally PM, Wallace MR, Gusella JF, Weder NS, Huntington disease: estimation of heterozygote status using linked genetic markers. Genet Epidemiol 1984:1:81 ~ 88.)。

有一个未受累的亲代，但这只是反映出这种特殊的显性遗传病发病年龄的延迟。应该注意的是发病年龄这样晚，致使致病基因在群体中传递，因为个体在尚未发病前，他们已生育孩子并把突变基因传给了下一代。上面的例子同时强调了"遗传的"和"先天的"(出生就存在的)不是同义的。显然，遗传病并不都是先天的。反之，先天性缺陷也不都是遗传的，如药物反应停所致的先天畸形是没有任何遗传基础的。

第四，偶尔会在一个早先没有该病家族史的家系中发现两个或更多的常染色体显性遗传病儿童。因为新的突变是罕见事件，所以不可能出现重复突变受累的个体。更为可能的解释是**生殖系嵌合体**(germline mosaicism)，即在胚胎生命过程中突变影响了配子前体，导致个体的某些或全部配子携有突变，而他(或她)的受累的体细胞数目不足以引起任何临床症状。像这样临床上未受累的个体可能有多个受累的子代。生殖系嵌合体已在产前致死性的成骨不全(见第七章)和神经纤维瘤以及 X 连锁隐性遗传病如 Duchenne 肌营养不良、甲型血友病的 DNA 分析中得到证实。

最后，在有些病例中推定的父亲并非是受累患儿的真正生物学父亲，这是常染色体显性疾病出现明显的散发病例的第五种原因。在一次大规模调查中，这种情况的频率约占 5% ~ 10%。

(二) 遗传异质性

大多数人类遗传病表现出**遗传异质性**(genetic heterogeneity)。遗传异质性就是不同的突变可引起相同的或相似的表型。我们进一步区别**等位基因异质性**(allelic heterogeneity)和**基因座异质性**(locus heterogeneity)。等位基因异质性是指在同一基因座上不同的突变,而基因座异质性是指在不同基因座上的突变。成骨不全(OI)的症状是骨脆弱,出生前偶然会自发骨折,或在生活中受到轻微的外伤后发生骨折。最近几种 OI 综合征分子基础的研究进展已表明这是等位基因异质性和基因座异质性的明显例子。位于 7 号染色体上编码 1 型胶原 α1 链的基因或位于 17 号染色体上编码 α2 链基因的突变均可改变胶原的结构而引起 OI。这就是基因座异质性。而且,如在第七章中讨论的那样,编码 α1 链的 *COL1A1* 基因的不同突变常能引起严重程度不同的 OI,这就称为等位基因异质性。基因座异质性的其他例子还包括第十二章讨论的遗传性结肠癌综合征。等位基因异质性实际上是人类孟德尔疾病的规律(软骨发育不全和镰状细胞贫血除外),例子包括 β 地贫(第六章),葡萄糖-6-磷酸脱氢酶缺乏症(第七章)和家族性高胆固醇血症(第七章)。

(三) 表现度不一致

突变的基因具有多效性,即可影响若干器官系统及其功能,经常呈现**表现度不一致**(variable expressivity)。虽然一个性状的表现度常与外显率混淆,但它是指表型的性质和严重程度。表现度不一致是常染色体显性性状的一种常见特征。Marfan 综合征就是一个好例子(图 3-8)。该病是由于肌原纤维蛋白基因突变引起的,肌原纤维蛋白是细胞外微纤维的主要成分,它影响身体的结缔组织,主要是骨骼系统(细长肢端,关节松弛,脊柱、胸骨畸形)、眼睛(近视眼和晶体错位)和心脏(瓣膜无力,主动脉根部膨大,有时主动脉破裂和突然死亡)。患 Marfan 综合征的个体可能仅累及两个或所有的三个主要系统,其表现的严重程度可能非常广泛。而且如图 3-9 系谱所示,在同一个家系中可能携有相同的突变等位基因的受累个体间,发病情况却不一致。对该家系的 4 个患者和 5 个未受累成员进行 DNA 分析,提供了特殊的信息。研究证实在所有 4 个接受测试的受累个体均发现相同突变。尽管如此,受累家系成员临床症状发作的年龄、累及器官系统的严重程度和分布是完全不同的。这些研究表明,Marfan 综合征表型不单纯是肌原纤维蛋白基因座突变的基因型所决定。表

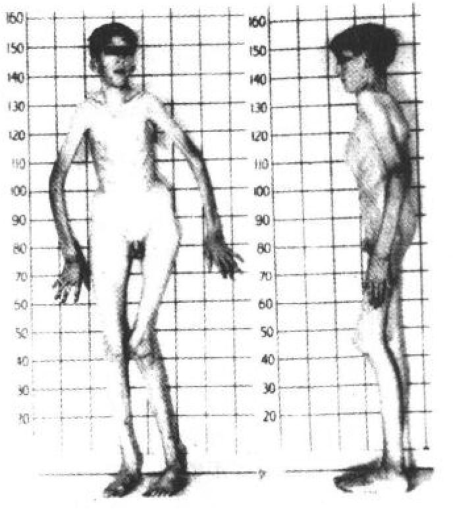


图 3-8 Marfan 综合征 17 岁男孩。肢端细长,细蜘蛛趾状手指,扁平足,胸骨、脊柱畸形(引自 Mckasick VA. Heritable disorders of connective tissue. 4 th ed. St. Louis: CV Mosby, 1972:67.)。

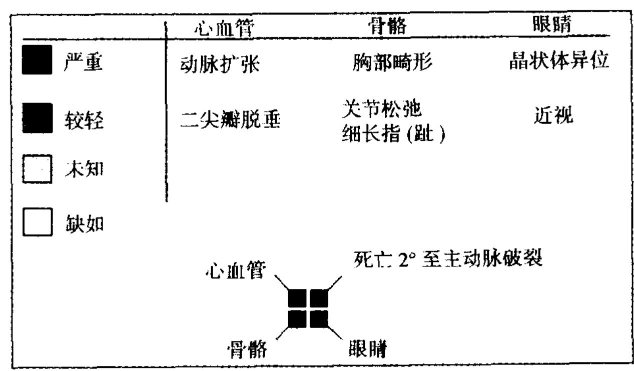


图 3-9 常染色体显性遗传 Marfan 综合征的系谱，说明表现度不一致。个体Ⅳ-1,2 和 3,Ⅴ-1,2,3,4 和 5,Ⅵ-1 进行 DNA 分析。所有受累的个体存在相同的肌原纤维蛋白基因的突变，所有受检家系成员均具有 Marfan 综合征的表现，至少累及一个器官系统。然而，在发作年龄、临床症状及受累器官分布和严重程度上存在明显的变异。Ⅰ-2、Ⅱ-2 和Ⅲ-3 由于严重心脏受累，年轻时死亡，并具有轻度眼部或骨骼受累。Ⅳ-4 和Ⅴ-3 有严重骨骼受累，带有轻度的或无眼部和心脏受累（引自 Dietz HC. Pyeritz RE. Puffenberger EG, et al. Marfan phenotype varibility in a family segrgating a missense mutation in the epidermal growth factor-like motif of the fibrillin gene. J Clin Invest 1992;89: 1674 ~ 1680.）。

现度不一致必定受环境的影响和（或）其他基因的作用，这些基因修饰 Marfan 综合征突变基因的表达。事实上，第Ⅴ代中受累的单卵双生子相同的表型说明其他基因座的等位基因影响突变的肌原纤维蛋白等位基因的表达。

表现度不一致更明显的例子是Ⅰ型神经纤维瘤(NFⅠ)或雷克林豪森病(Von Recklinghausen disease)(图 3-10)。这种常染色体显性遗传病的特征是皮肤上有棕色斑点,称

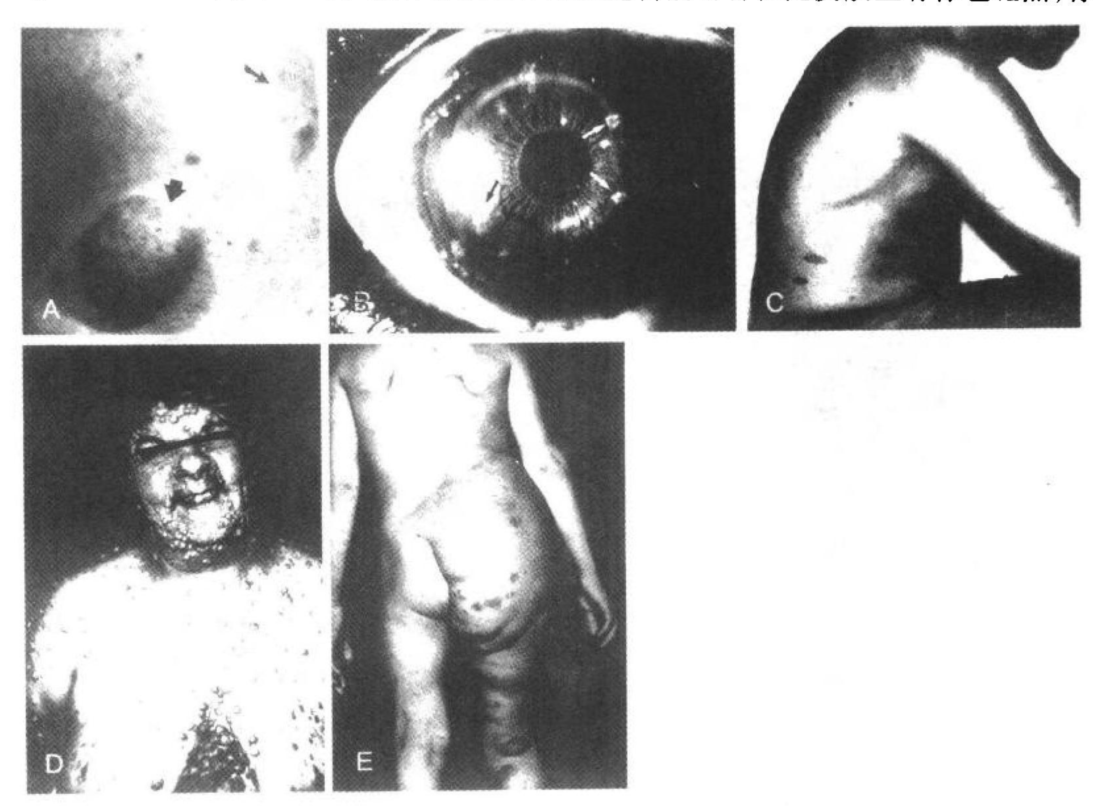


图 3-10 Ⅰ型神经纤维瘤的各种表现。A. 典型的牛奶咖啡斑和皮肤神经纤维瘤;B. 错钩小结;C. 在一个年轻男性躯干上具有牛奶咖啡斑,但没有观察到神经纤维瘤特征;D. 多个神经纤维瘤女性;E. 10 岁男孩,半边臀部大块过度生长,继发丛状神经纤维瘤,累及骨盆和周围软组织(引自 Crowe FW, Schull WJ, Neel JV. A clinical pathological and genetic study of multiple neurofibro-matosis. Springfiled, IL: Charles C Thomas, 1956:34.)。

"牛奶咖啡斑"、眼部虹膜上有良性错钩小结(lisch nodules)、表皮和皮下神经纤维瘤、肉质样良性肿瘤常常沿神经生长。神经纤维瘤的特征通常是在青春期出现,并可在一生中增加到几千个。此外,该病有各种表现,几乎身体的每个系统都受影响。高度脉管丛状的神经纤维瘤浸润周围组织,可能引起严重的生长异常和畸形。这种疾病的表现范围是从仅出现牛奶咖啡斑和错钩小结到损伤面容的表皮神经纤维瘤,以及严重畸形和功能损伤,很少有病人幸免。所有受累的个体都在相同神经纤维瘤基因座有突变。因此,家族内的不一致并非由于基因座异质性。尚无方法预测任何一个受累个体的表现将如何严重,也无任何证据表明在一个既定的家系中"真实繁育"(breed true)受累的严重程度。

(四) 遗传不稳定性和早现遗传

经过几十年的观察发现,一些显性遗传病表现出在连续世代中发病年龄一代比一代

早,且病情也加重。这种现象称为**早现遗传**(anticipation),最明显的例子是强直性肌营养不良。这是最常见的累及成人的肌营养不良。强直性肌营养不良的主要特征是肌无力,从面部开始(发生面罩样脸和上睑下垂)(图 3-11),然后颈、手逐渐遍及全身,从肌强直或肌无力到肌肉收缩松弛。强直性肌营养不良也影响心肌和平滑肌,与早期白内障、免疫球蛋白异常有关,并常有轻度智力低下。

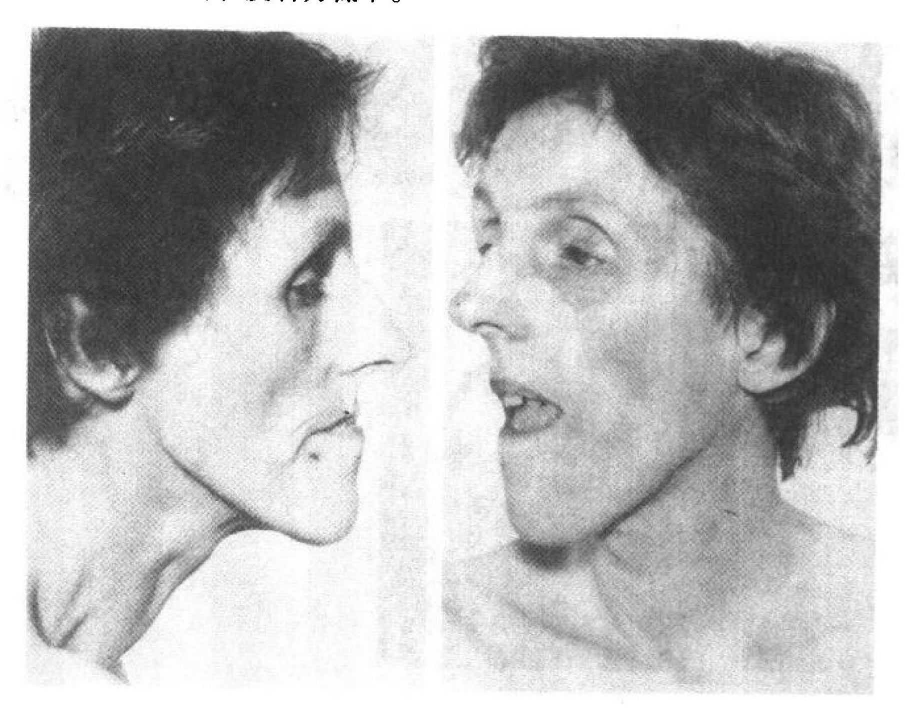


图 3-11 强直性肌营养不良成年病人。上睑下垂,面部消瘦,下颌和胸锁乳突肌消瘦。

在肌强直蛋白基因克隆以前,早现遗传没有生物学的解释,而且被认为反映了对临床受累患者人为的确认。然而,肌强直蛋白基因的分析表明,在该基因的 3′端非翻译区存在一个 GCT(或 CTG)三联体的多个重复。正常个体有 5 ~ 35 个三联体拷贝,而受累的个体则 > 50,有时达数百个到 > 1 000 个拷贝。重要的是该基因的这个区域是不稳定的,使其易感动态突变,表现为连续世代中重复数目的增加(图 3-12)。而且,肌强直性营养不良的严重程度和发病年龄与三联体重复数目密切相关(图 3-12)。罕见的先天性强直性肌营养不良婴儿的特征为严重张力减退(floppiness)和智力低下,而在这些婴儿中发现的重复数目最高。先天性类型只发生在婴儿,受累的母亲所生子女均有可能发生,而且在女性减数分裂中重复数目的增加最多。因此,对强直性肌营养不良建立长期的临床观察,使早现遗传现在有了遗传不稳定和三联体重复扩展的分子基础。这种遗传不稳定和性别特异性扩展引起先天性强直性肌营养不良的机制仍需确定。在脆性 X 染色体综合征(第八章中讨论)以及几种显性遗传的、成年发作神经退行性疾病包括 Huntington 病中均已发现了三联体重复扩展。在后一组病中有 10 ~ 30 个 CAG 三联体扩展,编码多聚谷氨酰胺域,在受累的个体中大约扩展 2 ~ 4 倍。发病年龄与重复数目的相关性虽然不像

强直性肌营养不良那么明显，但也观察到了早现遗传。

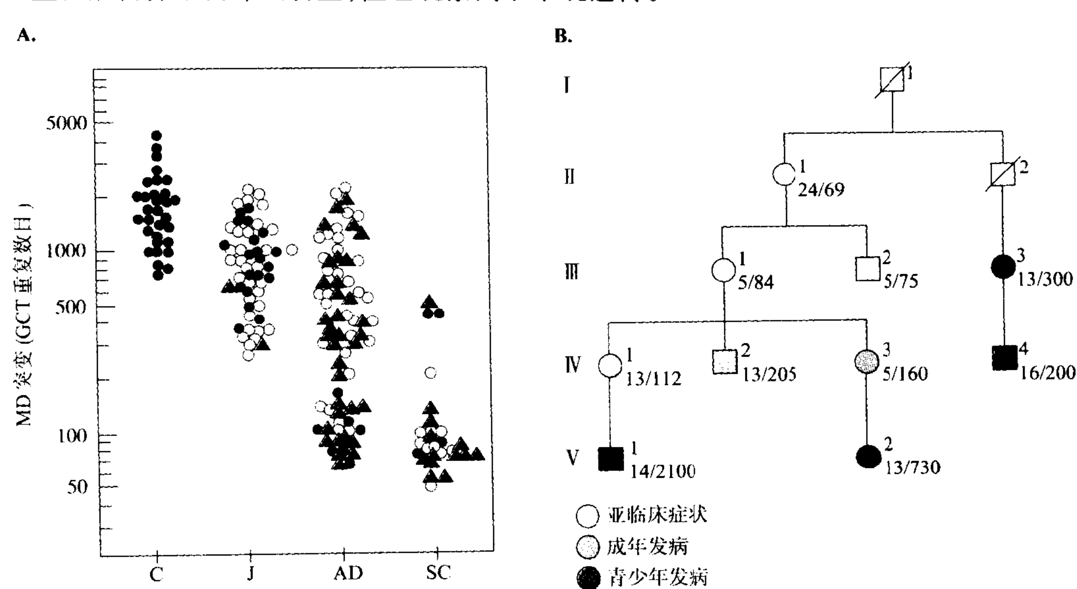


图 3-12 强直性肌营养不良(DM)发病年龄和/或严重程度与三核苷酸重复长度之间的关系。A.在197位患者中强直性肌营养不良GCT重复的大小和发病年龄。GCT重复扩展大小用重复数目来表达，与发病年龄和已知传递亲代的性别有关。C表示发病年龄范畴是从先天的症状至1岁；J表示年轻时发病，从1岁到小于20岁；AD表示成年，发病年龄20岁或更晚；SC表示无症状，区别于白内障和异常，但无肌电图诊断。传递的亲代性别注明●为母亲，○为父亲，▲为未知(引自 Redman JB, Fenwick RG, Fu Y-H, Pizzuti A, and Caskey CT. Relationship between parental trinucleotide GCT repeat length and severity of myotonic dystrophy in offsping JAMA 1993;269:1960～1965.)。B.为强直性肌营养不良早现遗传一个家系系谱。在中斜线左侧数字表示正常等位基因(三核苷酸重复数目)大小，右侧数字表明扩增的等位基因的大小(引自 Redmam JB, Fenwick RG, Fu Y-H, Pizzuti A, and Caskey CT. Relationship between parental trinucleotide GCT repeat length and severity of myotonic dystrophy in off spring JAMA 1993;269:1960～1965.)。

(五) 纯合性受累个体

对于很罕见的常染色体显性疾病受累的个体是杂合子。然而，当这个基因极为常见时，两个杂合子受累双亲婚配可产生纯合子受累后代。在大多数病例中，纯合子患者比杂合子患者严重的多。家族性高胆固醇血症就是很好的说明，该病纯合子胆固醇高水平增高，并在儿童中发生动脉粥样硬化心血管疾病。同样，纯合性软骨发育不全和马凡综合征在婴儿时致死。相反纯合子 Huntington 病并不比杂合子个体病情严重。

二、常染色体隐性遗传

常染色体隐性遗传病只有在突变基因的纯合性个体(或是在同一基因座上两个不同的突变等位基因的复合杂合子)的情况下才表现临床特征。其特征性的家系模式

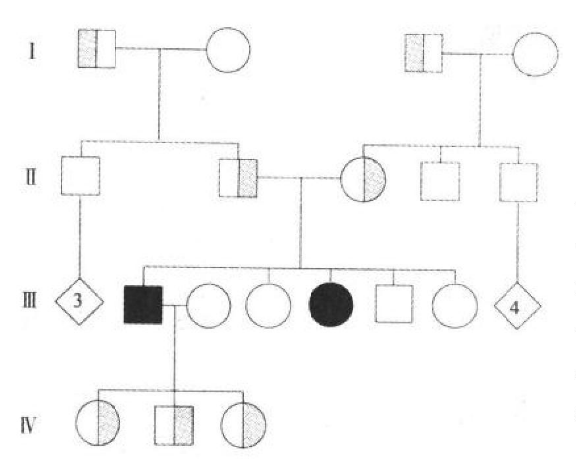


图 3-13 常染色体隐性性状系谱。受累个体(用实心黑色标志)仅在一代中出现。注意:受累儿童的双亲是肯定杂合子(用一半阴影灰色标志表明),他们都是纯合受累个体与正常个体婚配后产生的后代。

(图 3-13)是水平遗传而不是垂直的,受累个体仅限于单个同胞,而且在多个世代中不会出现这一疾病。男性和女性受累的概率相同。通常的婚配(图 3-14)是在两个临床表现正常却是突变等位基因杂合子的个体之间进行。每个后代将有 1/4 的机会是突变等位基因纯合性个体而受累;他(她)有 1/4 的机会是正常或野生型等位基因纯合性个体且临床表现正常;他(她)有 2/4 的机会是突变等位基因的杂合子,并与他(她)的每个双亲一样是临床正常的携带者。

两个具有相同突变等位基因的杂合性个体婚配的概率取决于杂合子在群体中的频率。反过来,这个概率又构成突变等位基因频率的函数。基因频率将在第四章进行更加详细的讨论。然而,很明显,如果突变等位基因在群体中是罕见的,那么任何两个携有相同突变等位基因的个体结婚的概率就会很低。倘若他们有亲缘关系并从共同祖先遗传相同的突变等位基因,那么这个概率就会增加。因此,近亲间的婚配,即**近亲婚配**(consanguinity)常见于子女患有罕见常染色体隐性遗传病的父母。另一方面,如果群体中突变基因很普遍,诸如白种人群的囊性纤维化,北欧犹太人(Ashkenazi European Jews)的 Tay-Sachs 病,或是非裔美国人(African-American)的镰状细胞贫血症,那么这种杂合子婚配更为常见,而亲代的近亲婚配将不像预期所见那么多。

由于若双亲二人都是携带者,其子女受累的风险仅为 1/4,又由于多数美国人家庭小(有 2 到 3 个孩子),因此,大多数常染色体隐性遗传病受累个体呈现散发病例,即在这个家庭中,

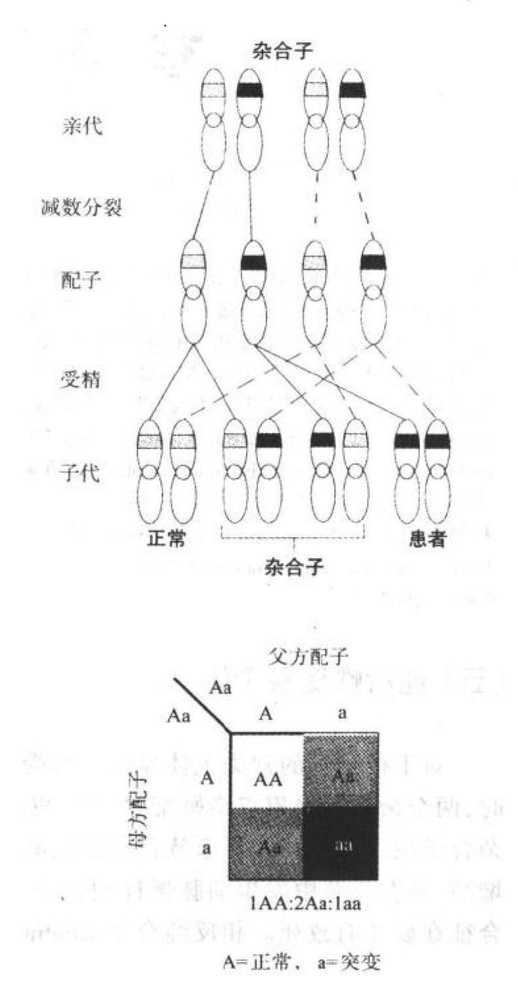


图 3-14 常染色体隐性遗传。婚配图和 Punnett 方框图表示临床正常却是突变等位基因杂合子的两个个体婚配后的结果。与图 3-4 一样,突变等位基因用黑色表示,受累个体用实心黑色方格表示,杂合携带者用灰色方格表示。

仅有一个受累成员。如果两个或更多的孩子患病，则这种病更可能被认为是遗传病。然而，应该注意的是大多数常染色体隐性遗传病将表现为散发病例，同时，重要的是由于缺乏明确的家族史而不应忽视他们的遗传病因学。

这个观点用图 3-15 说明。如果能确定所有 3 个孩子的家庭的父母双亲都是相同突变等位基因的携带者，那么在这些家庭中几乎一半（42%）将无受累儿童。因为事实上，只有存在受累儿童，我们才能识别杂合性携带者之间的婚配，所以，在没有进行杂合携带者筛查检测时这样的家庭是不能够确定的。这类 3 个孩子家庭中另有 42% 可能有单个受累儿童，而 14% 可能有 2 个受累儿童，仅有 2% 为全部 3 个孩子都受累。即使是在这些因至少有一个患病的子代而被确认的家庭中，几乎 3/4（27/37 或 73%）只有单个患儿，因此，这种疾病在这些家庭中呈散发性。

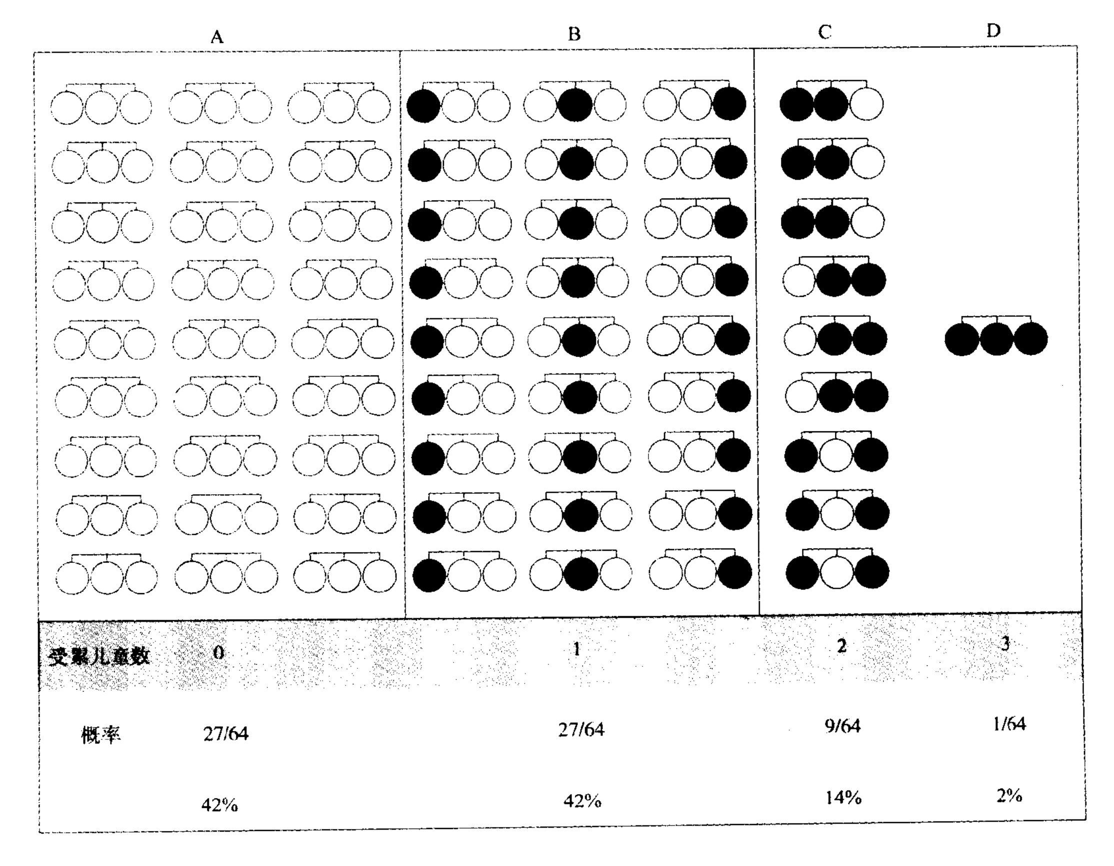


图 3-15 常染色体隐性性状杂合子婚配后出生的三个孩子中，受累 0、1、2 或 3 个的概率。在这个图解中，圆圈代表两种性别的后代，黑色标志表示受累个体，概率能够用二项式计算。注意：对于有三个孩子的家庭，几乎一半的婚配（42%）不会导致受累后代；而几乎一半将会导致单个受累后代（引自 Li CC. Human Genetics. New York: McGraw-Hill, 1961.）。

那么，我们怎样识别常染色体隐性性状呢？很明显，在一大群同胞中出现多个受累同胞而且亲代是近亲婚配时，则表明存在一种罕见的隐性疾病。然而，更有力的证据是在肯定杂合子中证明存在部分缺陷。Tay-Sachs 病（或 GM_2 神经节苷酯贮积症）是一种退

行性神经性疾病。在该病的受累儿童中实际上缺乏一种溶酶体酶即氨基己糖苷酯酶 A 的活性。Tay-Sachs 病患儿的父母临床表现正常，通常该酶具有大约正常数量的一半，这与通过家系分析说明此病为常染色体隐性性质一致。这种杂合子检测技术也能在频发携带者的群体中进行有效的筛查计划，正如 Tay-Sachs 病在北欧犹太人中大约占 3%，镰状细胞贫血在非裔美国人中占 8%。现在许多隐性遗传病分子方法的确立可以更精确地鉴定携带者。

镰状细胞贫血也证明显性或隐性是性状而不是基因。镰状细胞贫血(在第六章中详细讨论)是一种常染色体隐性遗传病，它只发生在突变 β^S 珠蛋白基因纯合性个体。然而，突变 β^S 基因的表达容易在杂合性携带者中检测到，即在具有镰状细胞症状而临床表现正常的个体中检测出突变 β^S 基因表达。因此，镰状细胞症状是一种常染色体显性性状。

引起常染色体隐性遗传病的突变等位基因纯合个体的所有后代都肯定是杂合性携带者，那么可以推断此婚配是发生在受累纯合子和正常纯合子之间(图 3-13 第Ⅳ代和

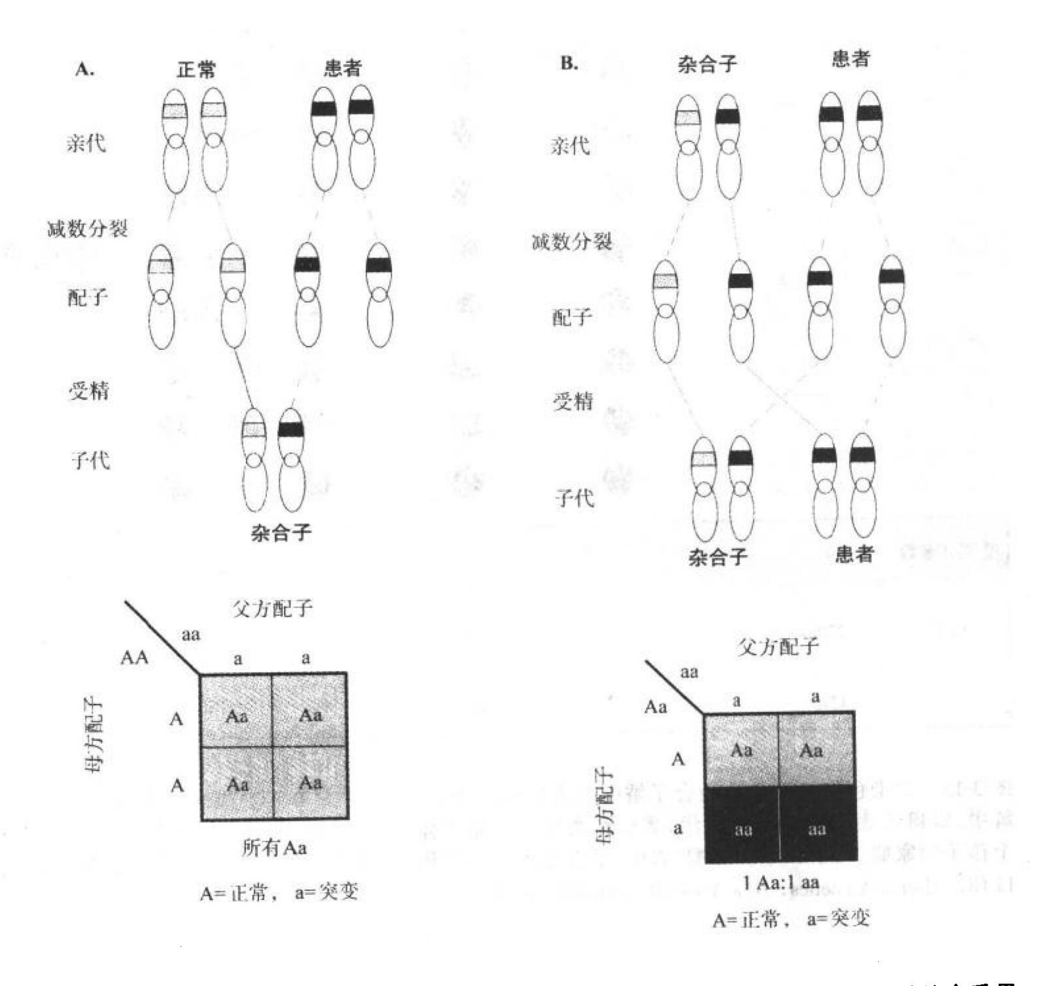


图 3-16 婚配图和 Punnett 方格图表明纯合受累个体与纯合正常个体婚配后的结果(A)以及纯合受累个体与临床正常的杂合个体婚配后的结果(B)。注意：A 中所有后代均是肯定杂合子，B 中每个后代有 50% 的概率是纯合性患者。

图 3-16A)。然而,当纯合受累个体与杂合携带者婚配时,其子代中将有 50%的概率也是纯合的突变等位基因,结果其遗传模式与常染色体显性相似,但认为是拟显性(图 3-16B)。这样的一种婚配不可能是偶然发生的,但是如果纯合受累个体与一近亲,例如一级表兄妹婚配,或者如果这是个普遍的突变等位基因,都会更有可能出现上述情况。

两个具有相同突变等位基因的纯合个体婚配,产生的所有子代都是纯合性受累个体。图 3-17 所示的系谱是后一类型婚配的实例,即涉及常染色体隐性先天耳聋个体的婚配。Ⅰ-3 和Ⅰ-4 以及Ⅱ-10 和Ⅱ-11 个体的所有子代同他们的父母一样都是先天性耳聋患者。然而,Ⅲ-7 和Ⅲ-9 个体的后代(Ⅳ-1 ~ Ⅳ-6)没有一个是耳聋患者。对这种奇异现象最简单的解释就是这些子代的父母(Ⅲ-7 和Ⅲ-9)是引起先天性耳聋的不同非等位(nonallelic)突变基因纯合性。因此,他们的六个子女都是双杂合子,即他们是引起耳聋的两个不同的非等位突变基因杂合子。因为对于每一个等位基因来说他们都不是纯合性,所以他们的听力是正常的。这个家系是基因座异质性的一个例子,也就是说在不同基因座的突变都会引起相同的表型即先天性耳聋。

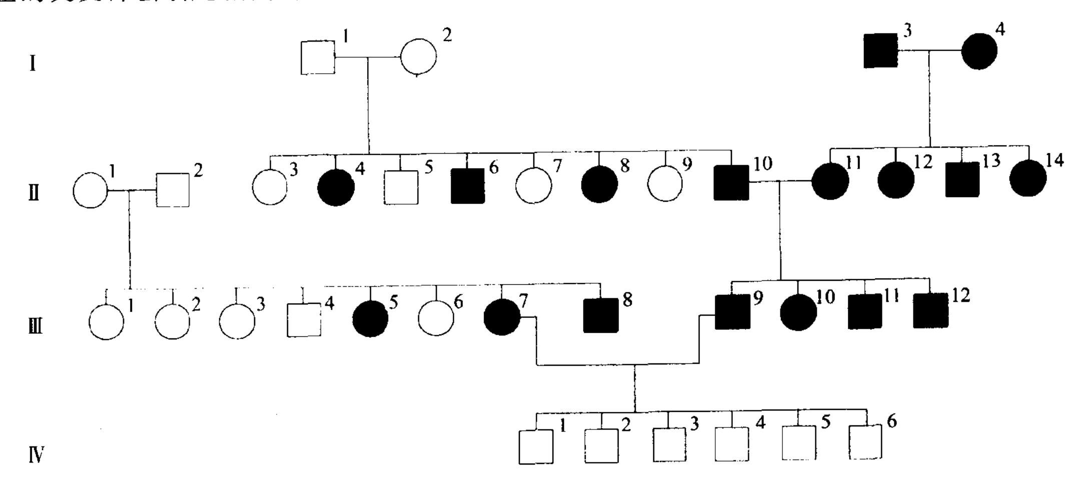


图 3-17 常染色体隐性先天耳聋家系的系谱。第Ⅳ代所有个体听力均正常,他们可能是两个非等位突变基因双杂合子(引自 Li CC. Human Genetics. New York:McGraw-Hill,1961.)。

三、X 连锁遗传

X 连锁疾病是由于 X 染色体上的突变基因所引起的。X 连锁突变基因在男性是完全表达的,因为男性只有一个 X 染色体,也就是说男性的 X 连锁基因是**半合性的**(hemizygous)。在人类女性所有体细胞都含有两条 X 染色体,然而只有一条 X 染色体具有遗传活性,两条 X 染色体之一在胚胎发育早期就随机并永久地失活,这一过程称作莱昂(Lyon)化(在第八章讨论)。由于 X 染色体随机的失活,所以对于 X 连锁突变基因杂合性的妇女来说,X 连锁性状的表达是可变的。由一个突变的 X 连锁基因所编码的一种疾病,对于一个杂合子女性来说,在临床上可能表达也可能不表达。杂合性女性的疾病在临床上几乎不表达称为 **X 连锁隐性**(X-linked recessive)。

典型的 X 连锁隐性遗传病是甲型血友病和 Duchenne 肌营养不良,前者是由于缺乏一

种重要的凝血因子即Ⅷ因子,以软组织出血和经常关节出血为主要特征。Duchenne肌营养不良是肌肉营养不良的一种类型,男孩不到5岁就开始发病,并发生进行性肌无力,通常在20多岁死于呼吸系统并发症。图3-18示一个典型的Duchenne肌营养不良患者家系,只有男性受累,但疾病是由杂合携带者的健康女性传递的,因此可以在家系母方的男性亲属中追寻该病的来源。

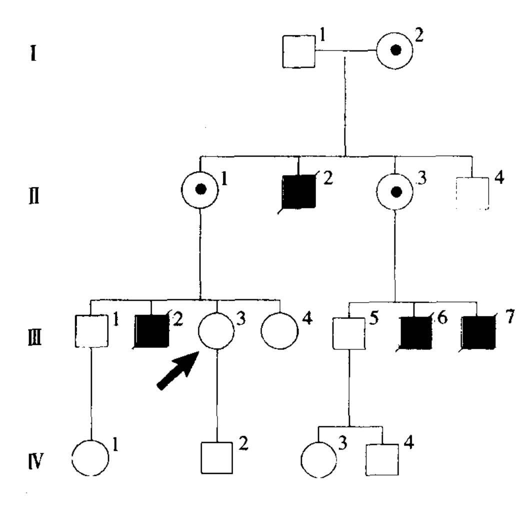


图3-18 X链锁隐性Duchenne肌营养不良家系的系谱。肯定杂合子用中间用有黑点的圆圈表示。

X连锁隐性性状遗传模式是很有特征的,所以古人就对其有所认识。犹太法典(Talmud)为血友病提供了参考,在患有这种出血病的家系中通过"割礼"提供了验证。另一个例子是18世纪Nova Scotia有关对饮水者诅咒的传说。在一个炎热的夏天,一个吉普赛人在一家农舍面前停下来为他口渴的儿子讨杯水喝,农舍的女主人拒绝给水并把他们打发走了,于是这个吉普赛人便诅咒她:"你的儿子将受到一种可怕的口渴的折磨,这个诅咒将会在你女儿的儿子身上再一次应验,而且代代如此。"于是,此病就这样传了下来。追溯早期此地区的居民所患的这种疾病(称为肾原性的多尿症)是由于肾脏对抗利尿激素不敏感,导致排出大量的稀释尿液,脱水,以致口渴。

注意X染色体并牢记这些疾病是X连锁而不是"性连锁"就能更好地理解X连锁遗传。通常的婚配发生于杂合携带者女性和正常男性之间(图3-19A)。他们的每一个儿子都有50%的机会得到存在于X染色体上的突变基因而受累,女儿也同样有50%的机会遗传到这个突变基因,但由于她从父亲那里遗传到一条正常的X染色体,所以她们无临床症状,然而她们有50%的机会是携带者。第二种典型的婚配(图3-19B)发生于男性患者和纯合正常女性之间,例如血友病。在这种婚配中,所有的儿子均正常,因为他们从患病的父亲处得到的是Y染色体而不是X染色体,他们从母亲处获得一条正常X染色体。然而,他们的女儿将全部都是肯定携带者,因为她们从父亲处获得一条带有突变基因的X染色体。因此,X连锁隐性性状从父亲传给儿子是不可能的;这种缺乏男性→男性的传递是X连锁遗传的主要特点。

图3-20所示家系患有一种遗传性慢性肾病——灶性肾小球硬化,医治这些病人的医生们认为他们是"性连锁遗传"。然而,从上述的结论表明,因为存在男性→男性传递,故这种疾病不是X连锁遗传。此家系显示的"性连锁遗传"可能是误导,在这里不适用。在这个家系中,性状可能是常染色体显性,可能是限性或从性,即性状只在男性中表达或较女性表达要多,尽管突变基因是在常染色体上而不是在X染色体上。一个普遍的从性常染色体显性性状是男子的秃发,它主要累及男性而且经常以男性→男性方式逐代传递。

当一种基因极其普遍时,X连锁性状也能出现表面上男性→男性传递。例如,在非

A.

女性杂合子　男性正常

亲代

减数分裂

配子

受精

子代

女性正常　男性正常　女性杂合子　男性患者

父方配子 X^AY；母方配子 X^AX^a

母方配子 \ 父方配子	X^A	Y
X^a	X^AX^a	X^aY
X^A	X^AX^A	X^AY

$1X^AX^A:1X^AX^a$

$1X^AY:1X^aY$

B.

女性正常　男性患者

亲代

减数分裂

配子

受精

子代

女性杂合子　男性正常

父方配子 X^aY；母方配子 X^AX^A

母方配子 \ 父方配子	X^a	Y
X^A	X^AX^a	X^AY
X^A	X^AX^a	X^AY

$1X^AX^a:1X^AY$

A=正常,a=突变

图 3-19　X 连锁隐性遗传。婚配图和 Punnett 方格图表明临床正常的突变等位基因(用深色条表示)杂合性女性和正常男性婚配结果(A)以及纯合正常女性和半合受累男性婚配结果(B)。Y 染色体用黑色表示。注意:X 染色体上的突变基因无男性→男性传递,但是受累男性的所有女儿均是肯定携带者。

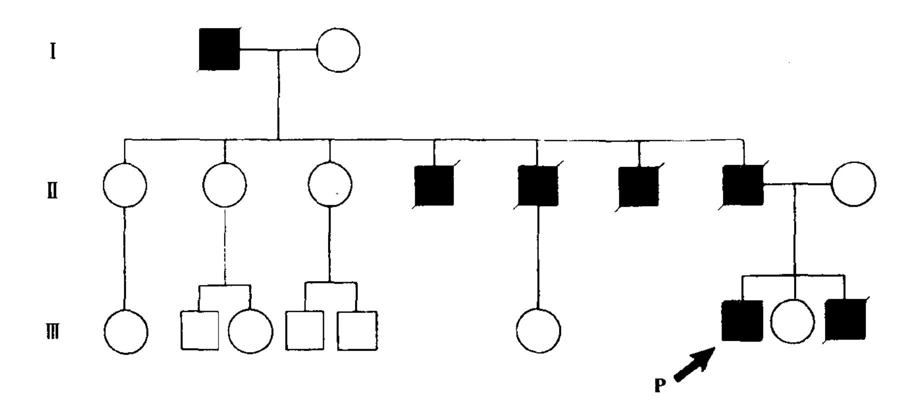


图 3-20　灶性肾小球硬化家系的系谱。注意:男性→男性传递规律与 X 连锁遗传不符。这个家系尽管只有男性受累,但它最符合常染色体显性遗传。

裔美国人和地中海地区的人群中普遍存在的葡萄糖-6-磷酸脱氢酶(G6PD)缺陷的基因(见第七章)。在非裔美国人中,大约18%的妇女是G6PD A⁻缺陷杂合携带者,大约10%男性是G6PD缺陷半合子。因此,如图3-21中所示G6PD缺陷表面上是从父亲(Ⅱ-4)传递给儿子(Ⅲ-1),事实上,X染色体上的基因是从杂合携带者母亲(Ⅱ-3)传递的,而与患病的父亲无关。值得注意的是,当X连锁基因以如此高的频率出现时,纯合受累女性也会以适度的频率出现。

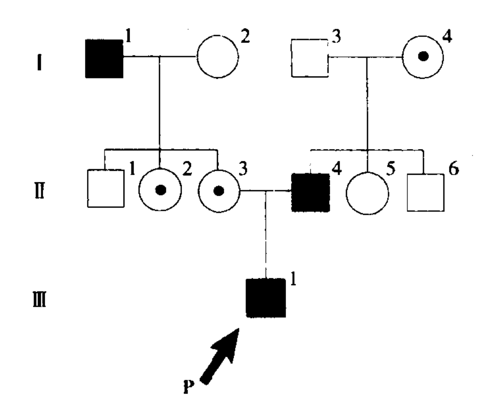


图3-21 X连锁隐性葡萄糖-6-磷酸脱氢酶缺陷家系的系谱。尽管表面存在男性→男性传递,但Ⅲ-1的突变等位基因一定获自他的肯定杂合子母亲(Ⅱ-3)。

在极少的情况下,会观察到妇女或女孩表达X连锁隐性遗传疾病性状。这有如下几方面原因。首先,如前所述突变等位基因极其普遍存在,结果观察到适度频率的纯合女性。例如,大约1%的非裔美国妇女存在G6PD(G6PD A⁻)缺陷,这些妇女是突变等位基因纯合性的,从患病的父亲处和携带者母亲各遗传一个突变基因拷贝。其次,女性也偶尔表达甚至像血友病那样罕见的X连锁基因,因为她是X染色体半合性的。这种情况也出现于Turner综合征,此时只有45条染色体而不是46条,而且仅有一条X染色体(见第八章)。第三,少见的染色体重排,即X染色体和常染色体上遗传物质互换(易位)导致了X染色体上遗传物质的丢失,并导致表达X连锁疾病。如女性DMD患者。发生的原因是正常的没有易位的X染色体优先失活(第八章),而有缺失基因的易位染色体在全身细胞中均具活性。这些事件的本质对于这种疾病的基因定位是至关重要的(见第九章)。第四,由于莱昂(Lyon)化(在第八章中讨论)的程度不同,杂合子妇女有可能表现出X连锁隐性疾病的临床症状。最后,一个女孩似乎患有X连锁隐性疾病如DMD,但事实上却是一种常染色体突变基因纯合性而导致了具有相似的临床表型。这是不同突变基因引起相似表型的基因座异质性的又一个例子。

图3-22表示**X连锁显性**(X-linked dominant)**性状**的一个典型家系。低磷酸盐(抗VD)佝偻病以低血磷、高尿磷、短身材和骨骼畸形为主要特征。这种遗传模式的一个主要特点也是无男性→男性传递,因为突变等位基因是在X染色体上。然而,在这种情况下女性也同样受累,而且男女患病比例大约是1:2,反映了男性、女性X染色体的比例。特别的是,在X连锁显性遗传病中,半合子男性患者的临床症状要比杂合子的女性患者恒定、严重得多,后者疾病的症状具有可变性。男性患者的女儿全都患病,因为她们从父亲处获得携有突变等位基因的X染色体。然而,男性患者的儿子全都正常,因为他们从父亲处得到的仅为Y染色体。女性患者的子女都有50%的机会获得带有突变等位基因的X染色体并因此患病。

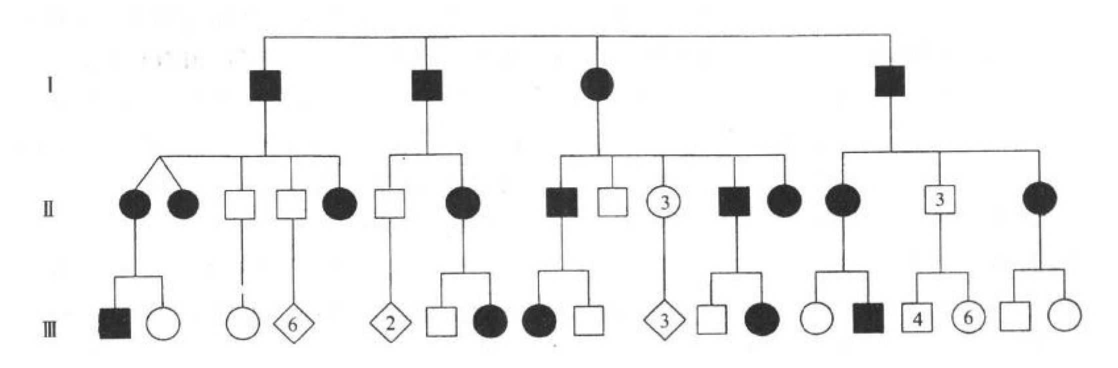
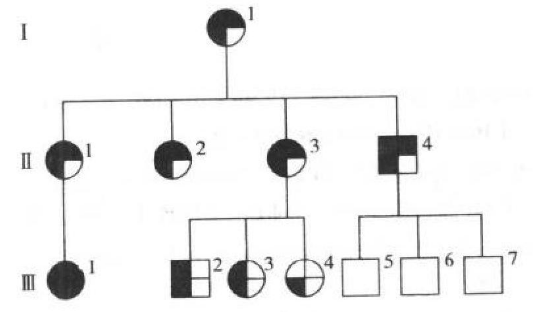


图 3-22 X 连锁显性低磷酸盐佝偻病家系的系谱。注意：男性患者所有女儿本身均患病。患病的男性和女性都能传递此病，但无男性→男性传递(引自 Williams TF, et al. In: The metabolic basis of inherited disease. Stanburt JB, Wyngaarden JB, Fredrickson DS, eds. New York: McGraw-Hill, 1960. Permission granted by McGraw-Hill.)。

四、线粒体遗传

孟德尔遗传描述的是细胞核内染色体编码基因的传递，然而，细胞质内的线粒体中也存在 DNA。每个细胞含有 2～100 个线粒体，而每个线粒体中有 5～10 个环状染色体。16.5kb 线粒体染色体编码 2 个 rRNA 基因，22 个 tRNA 基因和 13 个多肽基因，这些多肽构成氧化磷酸化相关酶的多个亚单位。线粒体 DNA 突变率大约是核 DNA 的 20 倍，在线粒体中可能再次产生的诱变氧化基团限制 DNA 修复能力。由于 mtDNA(线粒体 DNA)位于细胞质内，所以 mtDNA 遗传是专一的母系遗传(成熟的精子含有非常少的几乎没有线粒体的细胞质)。因此，只有女性能传递线粒体疾病，她们将突变传递给她们所有性别的后代(图 3-23)。因为 mtDNA 脱离核 DNA 自主复制，且在子细胞中线粒体分离不依赖于核染色体(这个过程称复制分离)，所以携有 mtDNA 突变的线粒体比例在体细胞和组织中有所不同。这种异质性称为**异序性**(heteroplasmy)，它在线粒体疾病的多样性和组织特异性表型中起到重要作用。此外，各种组织对氧化磷酸化的依赖性有所不同，如心脏、骨骼肌和中



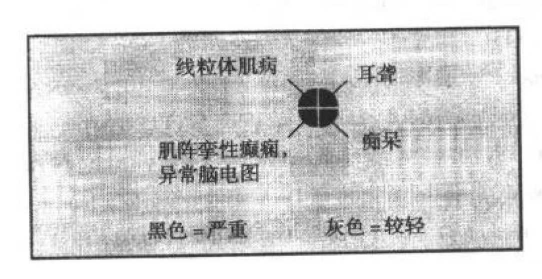


图 3-23 流行性红肌纤维线粒体脑肌病(MERRF)家系。注意：线粒体遗传的特点是母方传递给所有男性和女性后代而无男性传递。在受累成员中，临床严重程度与线粒体能量产生能力的缺陷有关，并且受累成员的变异性反映了线粒体突变的异序性(引自 Wallace DW, Zheng X, Lott MT, et al. Familial mitochondrial encephalomyoqathy (MERRF): genetic, pathophysiological, and bio-From Li CC. Human genetics. New York: McGraw-Hill, 1961. Chemical characterization of a mitochondrial DNA disease. Cell 1988;56: 601～610.)。

枢神经系统对氧化磷酸化的依赖性最强，因此线粒体病常以肌病和脑病为特征（分别是肌肉和大脑的疾病）。最后，氧化磷酸化作用随年龄增长而下降，或许与mtDNA突变的积累有关。因此，线粒体疾病的临床表型并非简单或直接与mtDNA基因型相关，而是若干因素的反应，包括核基因和线粒体基因所决定氧化磷酸化的遗传能力，体细胞中mtDNA突变的积累，异序性程度，氧化磷酸化的特异组织需要，以及年龄等。

最典型的线粒体病之一是Leber遗传性视神经病（LHON），该病丧失了中心视野，并继发视神经退化。通常在二十几岁时发生不明原因的突发性视力丧失，男性受累远较女性常见。该病呈母系遗传。编码呼吸链酶亚单位的三个不同线粒体基因发生了11种不同的错义突变，这方面已有描述。尽管表型都是相同的失明，但是失明的倾向以及起始的年龄都存在着很大的差异。其他线粒体疾病涉及到多种多样的肌肉和与其相关的神经系统的联合发病，而且在临床表型上也表现出高度的差异。

线粒体染色体上的一个非编码区（D环）具有高度的序列变异性（多态性，见第四章），已经在人类学和进化研究中用来追溯人类起源和亲缘关系。由于mtDNA多态性可以精确识别个体与他们母方之间的亲缘关系，因此它已用于解决独特的人权问题。mtDNA多态已用于为那些在阿根廷军事独裁时期父母"失踪"的孤儿们寻找他们的外祖母。

推 荐 读 物

Bennett RL, Steinhaus KA, Uhrich SB, et al. Recommendations for standardized human pedigree nomenclature. Am J Hum Genet 1995;56:745 ~ 752.

Dietz HC, Pyeritz RE. Review. Mutations in the human gene for fibrillin-1 (*FBN1*) in the Marfan syndrome and related disorders. Hum Mol Genet 1995;4:1799 ~ 1809.

Francomaho CA. Clinical implications of basic research. The genetic basis of dwarfism. N Engl J Med 1995;332:58 ~ 59.

Gusella JF, MacDonald ME. Huntington's disease and repeating trinucleotides. N Engl J Med 1994;330:1450 ~ 1451.

Harper PS, Harley HG, Reardon W, Shaw DJ. Review. Anticipation in myotonic dystrophy: new light on an old problem. Am J Hum Genet 1992;51:10 ~ 16.

Johns DR. Review. Mitochondrial DNA and disease. N Engl J Med 1995;333:638 ~ 644.

King MC. An application of DNA sequencing to a human rights problem. In: Friedmann T, ed. Molecular genetic medicine, vol. 1. Orlando: Academic Press, 1991:117 ~ 131.

MeKusick VA. Mendelian inheritance in man. 11th ed. Baltimore, Johns Hopkins University Press, 1994. *This valuable reference source is also available on the World Wide Web (or internet) at bttp://www. nebi. nlm. nib. gov/Omim/.*

Nawrotzki R, Blake DJ, Davies KE. Review. The genetic basis of neuromuscular disorders Trends Genet 1996;12:294 ~ 298.

Sutherland GR, Richards RI. Review. Simple tandem DNA repeats and human genetic disease. Proc Natal Acadia Sci USA. 1995;92: 3636 ~ 3641.

Warren ST. The expanding world of trinucleotide repeats. Science 1996;271:1374 ~ 1375.

第四章 群体遗传学和多因素遗传

几乎所有的人类疾病都具有家族性，即有患病亲属的个体的发病风险要高于无患病亲属的个体。

——J.H. Edwards(引自 Familial Predisposition In Man. Br Med Bull 1969;25:58 ~ 64)

在本章中，我们将探讨遗传学中两个需要较多定量方法的问题：群体遗传学和多因素遗传。群体遗传学是研究群体中基因的分布及逐代传递中维持和改变基因频率和基因型频率的因素。它是人类遗传学、进化研究和人类基因制图的中心部分。掌握群体遗传学对于法医中 DNA 检测的应用是必不可少的。在医学遗传学中，群体遗传学资料主要用于遗传咨询和制定遗传筛查项目。多因素遗传研究一类非单纯孟德尔方式遗传的，也与染色体异常无关的性状和疾病，但已有相当多的证据表明遗传因素在其成因中起重要作用。例子包括常见病如高血压和糖尿病，以及常见的出生缺陷如唇裂和腭裂。多因素模式为了解疾病的遗传易患性，并对患者及其家庭提供遗传咨询构建了框架。

一、群体遗传学

对于医学遗传学家来说，群体遗传学中最重要的概念是 Hardy-Weinberg 平衡。这是 1908 年英国数学家 G. H. Hardy 和德国内科医生 W. Weinberg 独自阐明的，它解释了为什么群体中显性性状不能自动地替代隐性性状。然而，它在医学遗传学中主要用于解释在随机婚配的大群体中，为什么等位基因频率在逐代传递中保持不变，以及在任何遗传基因座上，基因型频率是如何由该基因座等位基因的相对频率决定。

(一) Hardy-Weinberg 平衡

假设常染色体一个基因座上有两个等位基因(A 和 a)，在精子和卵细胞中的群体频率分别为 p = 等位基因 A 的频率和 q = 等位基因 a 的频率。因为只有两个等位基因，所以 $p + q = 1$。随机婚配(不考虑基因型的婚配)在数学意义上等同于随机混合，精子和卵细胞的结合情况如表 4-1 所示。每种基因型的频率列于括号内。因此，后代的基因型频率为：

基因型：	AA	Aa	aa
频　率：	p^2	$2pq$	q^2

在下一代中，父亲三种基因型的每一种可以与母亲三种基因型的每一种进行婚配，如表

4-2所示。每种婚配类型的频率列于括号内。每种婚配类型后代的基因型如表4-3所示。因此,经过连续几代后,每种基因型(AA,Aa和aa)的频率不变,则认为该群体处于Hardy-Weinberg平衡。

表4-1 后代的基因型频率

		父方配子	
		A(p)	a(q)
母方配子	A(p)	AA(p^2)	Aa(pq)
	a(q)	Aa(pq)	aa(q^2)

表4-2 婚配类型的频率

		父方基因型		
		AA(p^2)	Aa($2pq$)	aa(q^2)
母方基因型	AA(p^2)	AA × AA(p^4)	AA × Aa($2p^3q$)	AA × aa(p^2q^2)
	Aa($2pq$)	Aa × AA($2p^3q$)	Aa × Aa($4p^2q^2$)	Aa × aa($2pq^3$)
	aa(q^2)	aa × AA(p^2q^2)	aa × Aa($2pq^3$)	aa × aa(q^4)

表4-3 每种婚配类型后代中不同基因型的频率

婚配类型	频 率	后 代		
		AA	Aa	aa
AA × AA	p^4	p^4		
AA × Aa	$4p^3q$	$2p^3q$	$2p^3q$	
AA × aa	$2p^2q^2$		$2p^2q^2$	
Aa × Aa	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
Aa × aa	$4pq^3$		$2pq^3$	$2pq^3$
aa × aa	q^4			q^4

AA后代 $= p^4 + 2p^3q + p^2q^2 = p^2(p^2 + 2pq + q^2) = p^2(p + q)^2 = p^2(1)^2 = p^2$

Aa后代 $= 2p^3q + 4p^2q^2 + 2pq^3 = 2pq(p^2 + 2pq + q^2) = 2pq(p + q)^2 = 2pq(1)^2 = 2pq$

aa后代 $= p^2q^2 + 2pq^3 + q^4 = q^2(p^2 + 2pq + q^2) = q^2(p + q)^2 = q^2(1)^2 = q^2$

1. Hardy-Weinberg平衡的应用

Hardy-Weinberg平衡最重要的医学应用是在已知某一性状频率的群体中,确定等位基因频率和杂合子携带者频率。例如,在北欧和中欧地区,大约每2 000个白人中有1人患囊性纤维化。因此,q^2 或纯合受累个体的频率为1/2 000,则 q,即突变等位基因的频率为 $\sqrt{1/2\,000} = 1/45$ 或0.022。正常等位基因的频率为 $p = 44/45$ 或0.978。杂合子携带者的频率为 $2pq = 2 \times 44/45 \times 1/45 \approx 1/23$ 或0.043。因此,囊性纤维化等位基因在4%以上的

白种人中为杂合状态,这一事实对囊性纤维化家族的遗传咨询相当重要。

对于罕见的隐性遗传性状($q^2 \leqslant 0.0001$)来说,p 近似于 1,因此,杂合子携带者频率($2pq$)大约为 $2q$,即为突变等位基因频率的 2 倍。由此可见,对于任何罕见的常染色体隐性疾病来说,群体中杂合携带者的数量($2pq$)远远多于纯合受累个体的数量(q^2)。如表4-4所示,随着疾病发病率(q^2)的降低,这种比率(Aa/aa 或 $2pq/q^2$)升高。

表 4-4 基因型频率对杂合子/纯合子比率的影响

疾 病	发 病 率	等位基因频率		基因型频率			杂合子/纯合子
		A (p)	a (q)	AA (p^2)	Aa ($2pq$)	aa (q^2)	Aa/aa ($2pq/q^2$)
尿黑酸尿	1/1 000 000	0.999	0.001	0.998	0.002	0.000001	2 000
Tay-Sachs 病(非北欧犹太人)	1/360 000	0.998	0.0017	0.997	0.003	0.000003	1 071
Ⅰ型白化病(眼睛与皮肤)	1/40 000	0.995	0.005	0.990	0.010	0.0000025	400
苯丙酮尿症	1/10 000	0.990	0.010	0.980	0.020	0.0001	200
Tay-Sachs 病(北欧犹太人)	1/3 600	0.983	0.017	0.966	0.033	0.0003	110
囊性纤维化	1/2 000	0.978	0.022	0.956	0.044	0.0005	88
镰状细胞贫血	1/400	0.950	0.050	0.902	0.095	0.0025	38

由此可见,这种关系在制定遗传筛查项目时意义重大。例如:未建立筛查项目之前,在北欧犹太人(Ashkenazi Jewish)群体中,Tay-Sachs 病的发病率约为 1/3 600,即 0.0003。因此,突变等位基因频率(q)约为 0.017,杂合携带者频率($2pq$)为 0.033。北欧犹太人群杂合子间婚配的频率,即有生育患 Tay-Sachs 病子女风险的夫妇间婚配的频率,为 $2pq \times 2pq$,约为 1/1 000。假设每对夫妻有两个孩子,为避免一个 Tay-Sachs 病患儿的出生,必须在该群体中筛查约 2 000 对夫妻(并为这些夫妻提供产前诊断和选择性流产)。相比之下,在非北欧犹太人群体中,Tay-Sachs 病的发病率约为 1/360 000。因此,$q = 0.0017$,$2pq = 0.003$。非北欧犹太人群体 Tay-Sachs 病携带者间婚配的频率为 1/100 000,因此,为避免一个 Tay-Sachs 病患儿的出生,必须筛查 200 000 对夫妻。显然,对于这种罕见的疾病来说,筛查北欧犹太人群体比筛查整个群体更有效、更经济。事实上,筛查项目已使北美的北欧犹太人群体中 Tay-Sachs 病的发病率降低 90% 以上(见第十三章)。

对于 X 连锁基因来说,由于男性为 X 连锁基因的半合子,所以等位基因频率的估计与常染色体基因不同。因此,受累男性频率等于突变等位基因频率 q。对于一种相对罕见的 X 连锁隐性疾病而言,如甲型血友病(q,即发病率约为 1/5 000,$p \approx 1$),杂合携带者的频率仅是受累男性频率的 2 倍($2q = 1/2\ 500$)。然而,由于 $q \gg q^2$(血友病为 1/5 000 比 1/25 000 000),所以,受累男性与纯合受累女性的比率非常高。相反,对于 X 连锁显性疾病来说,受累男性约为受累女性的一半,$q:2q$。

(二) 影响基因频率的因素

Hardy-Weinberg 定律充其量是个估计值,因为只有在某些非常特定的条件下它才成立,而在人群中却难以满足这些条件。然而,值得医学遗传学家庆幸的是,这个定律的作用是使人们能够像上面描述的那样估计临床实用的等位基因频率。Hardy-Weinberg 平衡只有在一个随机婚配,没有选择、突变和迁移的很大的群体中才适用。这些条件的任何一种偏离都可能改变群体中等位基因频率,引起等位基因频率在逐代传递过程中升高或降低。

1. 非随机婚配

在人群中,婚配很少是随机的。无论一个子群体是否是依据种族、民族、宗教或者其他标准而确定,通常这个特定的子群体的成员很可能与该子群体内其他成员婚配。例如,先天性耳聋个体间婚配要比预计的随机婚配更常见。人群中非随机婚配的一种特殊形式为**近亲婚配**(consanguinity),即在近亲间的婚配。尽管近亲婚配不改变等位基因频率,但它的确可以降低杂合子数量而提高下一代中纯合子的比例,因此使不利的隐性表型面临选择。这种选择又可以在以后的各代中改变等位基因频率。近亲婚配不仅提高了后代中产生有害的隐性基因纯合子的风险,而且增加了后代对多基因或者多因素疾病和出生缺陷的易感性。正如本章后面所讨论的,患病风险与亲属级别成比例。近亲的程度可以用**亲缘系数**(coefficient of relationship, r)来表示。亲缘系数是指两个人从一个共同祖先遗传获得一个特定等位基因的概率。它也代表从一个共同祖先传递的、完全一致的所有基因的比例(表 4-5)。

表 4-5　不同亲属级别的亲缘系数

亲　属	级　别	亲缘系数
双亲 - 子女	一 级	1/2
同胞(兄弟姐妹)	一 级	1/2
叔 - 侄	二 级	1/4
一级表兄妹	三 级	1/8
二级表兄妹	五 级	1/32

正如第三章所述,近亲婚配提高了带有相同突变等位基因的两个杂合个体间婚配的概率。因此,可以推测出,在患有罕见常染色体隐性疾病的患儿的双亲中,近亲婚配的频率比较高。事实上也是如此,而且近亲婚配的频率与突变等位基因频率呈负相关。疾病越罕见,双亲近亲婚配的可能性越大。通过比较在北欧犹太人和非北欧犹太人群体中 Tay-Sachs 病的情况,可以阐明近亲婚配频率和隐性疾病发病率之间的关系。在非北欧犹太人群体中,$q = 0.0017$,双亲中近亲婚配的频率约为 4%,而在北欧犹太人群体中,$q = 0.017$,双亲中近亲婚配的频率低于 0.4%。在一般美国人群中一级表亲婚配的频率低于

0.1%；然而，在特定隔离群中，如哈特派信徒（Huutterites）或基督教门诺派中严紧派（old-order Amish），近亲婚配的频率相当高。

2. 小群体

由于政治、宗教或者地理上的原因，群体的一个小亚群可以在自然和/或社会上从该群体中分离出来，形成一个遗传隔离群。由于偶然因素，这个隔离群的建立者可能是某些隐性性状突变等位基因的携带者，以至于隔离群中这些等位基因的频率要高于整个群体。而且，在这个小群体中，实际的基因频率在逐代传递过程中差异很大，称为**遗传漂变**（genetic drift）。出于偶然，一等位基因可能不传递给下一代而从此消失，仅留下另一个等位基因，这种机制称为**建立者效应**（founder effect）。它可导致在一些遗传隔离群中某些罕见疾病的发病率高。例如，常染色体隐性 Ellis-van Creveld 综合征，其特征为短肢、侏儒、多指（趾）和先天性心脏病（图 4-1）。该综合征在宾夕法尼亚的基督教门诺派中的严紧派群体中等位基因频率 $q = 0.07$，而在一般群体中相当罕见。另一个可能的建立者效应的例子是在南非白人群体中发病率较高的斑驳卟啉病。这是一种常染色体显性疾病，其特征为急性腹痛发作，虚弱无力，阳光所致的皮肤病变等。人们认为该突变等位基因是 17 世纪末由荷兰移民引入到好望角部落的一个很小群体中，现在该病的发病率已达到1/400。北欧犹太人群体中某些遗传疾病的高发病率也是建立者效应的一种反映。



图 4-1 Ellis-van Creveld 综合征患儿（引自 McKusick VA, et al. Trans Assoc Am Physicians 1964;77:154.）。

3. 选择

选择反映环境因素对特定的表型，乃至基因型的作用，它可以是正性的，也可以是负性的。它是个体表型**生物适合度**（biological fitness, f）不同的结果。生物适合度用来衡量生育力大小，进而反映下一代基因库的分布情况。因此，选择可以在从妊娠至生殖期结束的任意时候发生。一个突变等位基因如果影响到生育力，即使它不引起疾病，也可能是"遗传致死性"的。

大多数有害的显性突变具有一个介于 0（遗传致死性）和 1（正常等位基因的适合度）之间的适合度。编码常染色体显性性状的突变等位基因在杂合子即可表达而面临直接选择。因此，选择压力的变化能够迅速地改变显性突变的等位基因频率。

相比之下，对编码常染色体隐性疾病的突变等位基因的选择进行得很慢。如上所述

(表 4-4),罕见的隐性疾病杂合携带者的频率远高于受累纯合子的频率。因此,大多数突变的等位基因由杂合子携带而不是由纯合子携带。对于一种发病率约为 1/15 000 的致死性常染色体隐性疾病而言,受累纯合子的生物适合度由 0 恢复到 1,只有经过 50 代后,才引起基因频率增加 1 倍。相反,临床表型正常的杂合携带者适合度的轻微增加可能极大地影响突变等位基因频率。至于 X 连锁隐性疾病如 Duchenne 肌营养不良和甲型血友病,群体中 1/3 的突变等位基因存在于半合子男性中而面临直接选择。提高受累男性的生物适合度,如改善甲型血友病的治疗,将会明显地增加突变等位基因的频率。

4. 突变

突变是指遗传物质发生的改变。这种变化的频率称为**突变率**(mutation rate, μ),一般用每代每个配子中每个基因座的突变数来表示。大多数人类突变率的资料来自于对罕见的常染色体显性性状的研究,因为它比隐性性状更容易估计突变率。对于罕见的常染色体显性性状而言,突变率 $\mu = n/2N$,其中 n 为双亲正常的受累患者数,N 为出生总数。因为常染色体基因座上任何一个等位基因的突变都可引起突变表型,所以分母为 $2N$ 而不是 N。基因座异质性(不同遗传基因座的突变导致相似的表型)和受累双亲外显率降低均可引起对突变率错误地过高估计。

各种常染色体显性疾病突变率的估计值从 10^{-6} 至高达 10^{-4} 变化不等。对于大多数人类基因而言,突变率的最佳估计值介于 $10^{-6} \sim 10^{-5}$ 基因座$^{-1}$·配子$^{-1}$·代$^{-1}$。有证据表明,随着父亲年龄的增加,某些常染色体显性性状的突变率增加。

两种重要的遗传疾病因其极高的突变率而值得注意。一个是 I 型神经纤维瘤(NF1),估计突变率为 10^{-4}。以前曾认为这是由于几个不同的遗传基因座都可以引起 NF1 所造成的,但现在这种解释已被否认(见第九章)。Duchenne 肌营养不良的突变率也非常高,约为 $0.2 \sim 1.0 \times 10^{-4}$。其高突变率可能是由于该基因非常大造成的(> 2400kb)。该基因是目前已知的、最大的人类基因,它可能成为公认的诱变剂作用的一个大"目标"(target)。

当处于平衡状态时,一种适合度降低的、罕见的常染色体显性性状如软骨发育不全,通过突变引入新的突变等位基因,又通过负性选择排除这些突变等位基因,以此来维持平衡。生物适合度降低得越严重,由新突变致病的患者比例越大。如致死性成骨不全Ⅱ型,该病所有的病例均由新突变产生。相比之下,Huntington 病的临床发病时间晚,生物适合度仅轻微受损,因此,在该病的所有病例中仅有很少一部分是由新突变造成的。对于遗传致死性 X 连锁隐性疾病如 Duchenne 肌营养不良(DMD)而言,存在于受累男性半合子中的所有基因占群体中突变等位基因的 1/3,这些突变基因经过一代后就消失了。根据 Haldane 假说,从疾病发病率的角度考虑,如果群体处于平衡状态,则新突变的数目必须等于消失的等位基因数目。因此,1/3 DMD 受累男性为新突变产生的。

可能引起人类突变的环境因素包括离子辐射和化学诱变剂等。尽管在其他哺乳动物和低等生物中已获得了较好的、关于辐射剂量与突变率关系的实验证据,但在人类群体中有关数据甚少。这方面的资料主要来源于对核爆炸或者辐射意外引起的辐射暴露的遗传效应的长期研究。已知许多天然的和人工合成的化学物质对其他生物系统有致

突变作用,它们对人类可能的致突变风险已成为公共健康关注的主要问题。

5. 迁移和基因流

群体迁入新的地区,并与当地的群体婚配,可以引起两个群体等位基因频率的变化,这种现象称为**基因流**(gene flow)。它可以解释B型血等位基因在亚洲高频率,而当向西迁移跨越欧洲时逐步降低的现象。

(三)多态性

到目前为止,主要讨论的是在群体中的大多数为单一正常(或野生型)等位基因的遗传基因座上罕见的突变等位基因。然而,在同一群体中,人类的许多基因座上有两个或两个以上的等位基因,以相当可观的频率出现。**多态性**(polymorphism)的概念为在一个群体中,存在两种或两种以上的、由遗传决定的不同表型,其中最罕见表型的频率不能单独依靠再发突变维持。事实上,如果一个或多个罕见等位基因的频率至少为0.01,携带该等位基因的杂合子频率大于2%,则认为该遗传基因座具有多态性。在已研究过的、编码蛋白质的人类遗传其因座中,1/3以上的遗传基因座具有多态性。例如:编码常见血型如ABO、MN、Rh系统的基因座,编码各种红细胞酶类和血清蛋白的基因座,编码主要组织相容性复合物细胞抗原的基因座。多态性可以作为遗传多样性的标志,是人类基因组制图的重要工具,这一点将在第九章详细讨论。涉及到限制酶识别位点或DNA短串联重复序列的DNA多态(第五章讨论)在这方面更具有特殊价值。

人类基因组的广泛多态性要考虑到不同基因座等位基因的多种组合,进而导致群体的高度遗传多样性和个体的遗传独特性。这种遗传多样性可解释各种疾病遗传易感性的差异,如本章后面提及的糖尿病。最好的研究范例之一是主要组织相容性复合物,即MHC(major histocompatibility complex)。它是一簇高度多态的基因,在免疫应答调控过程中发挥重要作用。该区域三类基因中的两大类,Ⅰ类(HLA-A,HLA-B和HLA-C)和Ⅱ类(DP,DQ和DR)又称人类白细胞抗原(human leukocyte antigen, HLA)基因,编码全部在抗原递呈给免疫T淋巴细胞时所必需的膜蛋白。因为MHC基因座紧密连锁,位于6号染色体不足4 000kb的区域,所以,在一特定染色体上的HLA等位基因通常作为单体型(haplotype)一起传递而表现出明显的连锁不平衡(在第九章讨论)。某些HLA抗原与特定的疾病显著相关,通常为自身免疫性疾病。这类疾病主要表现为机体对自身的组织产生免疫反应。最典型的例子是HLA-B27与强直性脊椎炎相关。强直性脊椎炎是一种慢性、炎性关节炎,主要侵袭脊柱和骶髂关节。仅有7%白种人是HLA-B27(+),而强直性脊椎炎患者95%为阳性。尽管不是所有B27(+)的个体都将患病,但是他们的患病风险比B27(-)个体高90倍。

一般认为,尽管在多数情况下,引起所谓平衡多态的选择压力难以估计,但是大多数涉及表达基因的、稳定的多态是对有利的突变等位基因选择的结果。选择导致平衡多态的最典型的例子为镰状细胞等位基因的平衡多态。该突变等位基因为纯合的个体表现为镰状细胞贫血,这是一种严重的溶血性贫血,通常在生育期前死亡。然而,在美国黑人

中,出生时患有镰状细胞贫血的约为1/400,在赤道非洲部分地区,该病的发病率达到1/25。因为受累纯合子的适合度明显降低,所以突变等位基因的高频率必定是由于该等位基因的杂合个体适合度增加造成的。在过去恶性疟疾(因感染镰状疟原虫而致病)常见的地区,镰状细胞贫血发病率也较高(图4-2)。在这些地区,许多人患有疟疾,生育力下降。镰状细胞等位基因的杂合携带者(具有镰状细胞性状的个体)对感染疟疾有较强的抵抗性,原因在于当红细胞中有镰状疟原虫寄生后,他们比正常无镰状细胞性状的个体更有效地从循环系统中清除红细胞。因为常染色体隐性性状的杂合子数目远多于受累纯合子数目(表4-4),所以,杂合子相对较小的选择优势可以超过纯合子较大的选择劣势。对镰状细胞贫血而言,具有镰状细胞性状的个体适合度增加20%,可以平衡镰状细胞贫血患者适合度降低85%。由此可以推测,在世界的一些地区控制疟疾的发生后,该等位基因频率应该降低。第六章讨论的一组常染色体隐性贫血——地中海贫血和第七章讨论的G6PD缺乏症与镰状细胞贫血有相同的地理分布,抗疟性可能是这些疾病发病率高的正性选择压力。

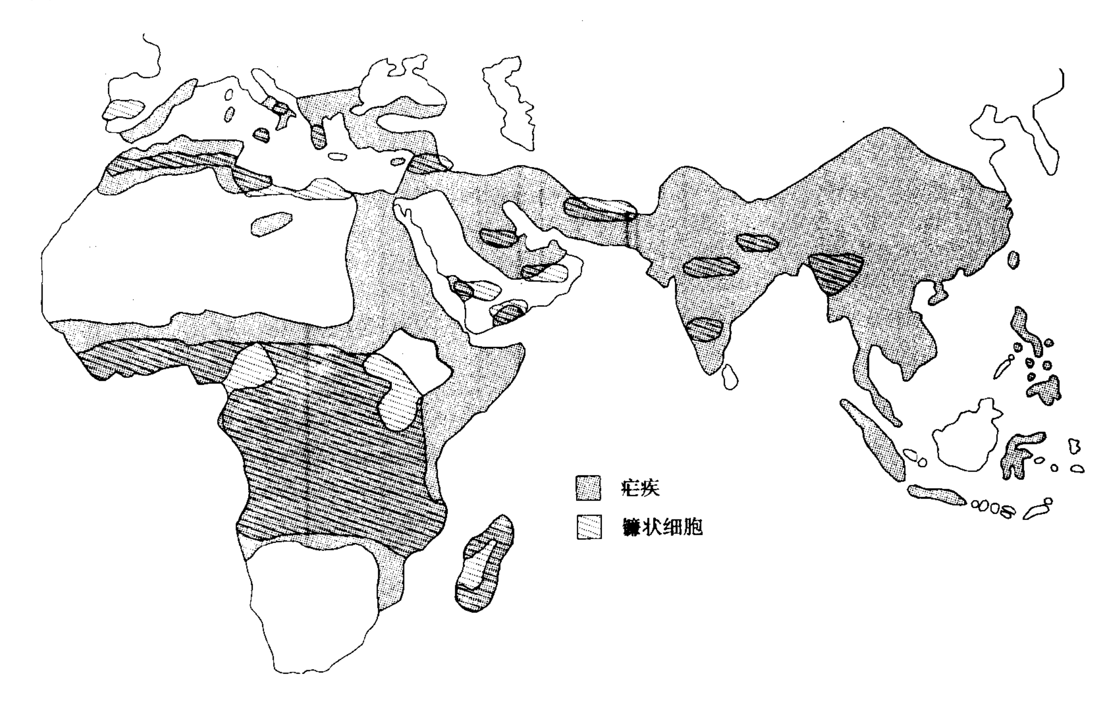


图4-2 1930年前恶性疟疾和镰状细胞贫血症的地理分布(引自Friedman MJ, Trager W. The biochemistry of resistance to Malaria. Sci Am 1981;244: 154~164.)。

二、多因素遗传和数量性状

尽管许多人类疾病呈单纯孟德尔式遗传或者与染色体异常有关,但大多数成年人常见病(如糖尿病、高血压、精神分裂症)和多数常见先天畸形(如唇裂、腭裂、神经管缺陷)却例外。然而,如下讨论,已有相当多的证据表明遗传因素在其病因学中的重要性。这种疾病的遗传易感性被认为是几个或多个基因座遗传变异的累积效应,每个基因座对表

型的影响相对较小。我们把“**多基因的**”(polygenic)确定为一类由许多不同基因作用引起的性状或疾病，每一个基因对表型仅产生有限的个别影响。把“**多因素的**”(multifactorial)确定为一类由多个环境因素和多个基因相互作用而产生的性状或疾病。事实上，这两个概念经常相互替换使用。多基因性状通常为数量性状而不是质量性状，在群体中连续分布，常为近似正态分布(Gaussian 分布)，例如身高和血压。

为了理解数量性状的正态分布，假设两个独立的遗传基因座 A 和 B，均影响收缩压。每个基因座有两个等位基因 A 和 a，B 和 b，在群体中频率相等 $p = q = 0.5$。假设群体处于 Hardy-Weinberg 平衡，则每个基因座的三种基因型(AA、Aa、aa 和 BB、Bb、bb)的比率为 $p^2:2pq:q^2 = (1/4):(1/2):(1/4)$。如表 4-6 所示，由二个基因座的基因型频率等于单个基因座基因型频率的乘积(独立性原理)。对感兴趣的假定的数量性状(1.33×10^4 Pa 基础水平之上的收缩压)用大写字母(A 或 B)标记表示等位基因使血压增加一个单位，平均加 1.33×10^3 Pa 血压，而用小写字母(a 或者 b)标记表示等位基因不改变血压，为 0 单位(表 4-6)。则如图 4-3 所示，该性状的群体分布情况近似于正态分布。

表 4-6 两个基因座位点两个等位基因模式决定的收缩压的频率分布*

	AA 1/4	Aa 1/2	aa 1/4
BB 1/4	1/16 ($40 \times 1.33 \times 10^2$)	2/16 ($30 \times 1.33 \times 10^2$)	1/16 ($20 \times 1.33 \times 10^2$)
Bb 1/2	2/16 ($30 \times 1.33 \times 10^2$)	4/16 ($20 \times 1.33 \times 10^2$)	1/16 ($10 \times 1.33 \times 10^2$)
bb 1/4	1/16 ($20 \times 1.33 \times 10^2$)	2/16 ($10 \times 1.33 \times 10^2$)	1/16 (0)

* 括号内的数值表示相对于基础血压 1.33×10^4 Pa 而言，每种基因型引起的收缩压的增高值(Pa)。

与性状有关的遗传基因座的数目越大，和/或该基因座的多态性越多(即在任一特定基因座，不同的等位基因数目越多)，则性状分布越有可能是一条平滑的正态曲线。环境因素的作用可能进一步修饰频率分布的形状。

实际上，正如上面假设的例子一样，数量性状明显连续的单峰分布代表了几种独立表型的总和。这一点通过红细胞酸性磷酸酶活性的研究得以说明，如图 4-4 所示。英国人群中有 6 种不同表型，区别在于酶电泳迁移率的不同，是由一个常染色体基因座的三个等位基因决定的。六种表型包括三种纯合基因型 AA、BB 和 CC 以及三种杂合基因型 BA、CA 和 CB。每种类型与酶活性的某一特征性范围相关。尽管具有同一酸性磷酸酶基因型的个体之间磷酸酶活性有相当大的差异，但是不同基因型间酶活性平均值的差异非常明显。图 4-4 显示 5 种常见

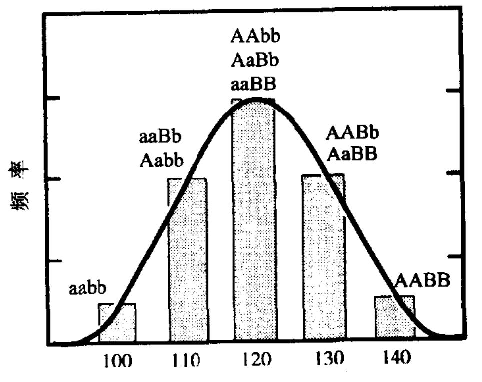


图 4-3 两个基因座两个等位基因模式决定的收缩压的频率分布。

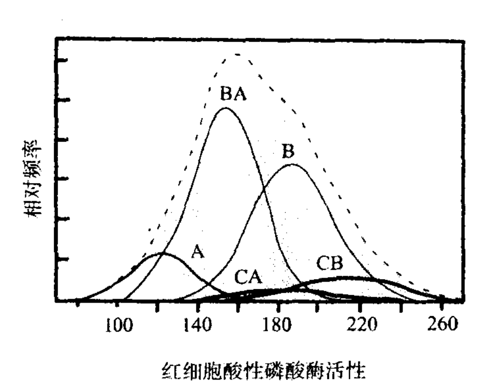


图 4-4 群体(虚线)及各种表型个体红细胞酸性磷酸酶活性的分布。实线是根据英国人群中不同表型的资料绘制(引自 Harris H. The principles of human biochemical genetics. 3rd ed Amsterdam: Elsevier/North-Holland, 1980:186.)。

的酸性磷酸酶基因型(CC 相对罕见)酶活性的个体分布情况(实线)和英国人群中酸性磷酸酶活性的总体分布情况(虚线)。很明显,整个群体的单峰分布是由 5 种独立的但又相互重叠的分布组成。

多基因和亲属关系

除了单卵双生子外,虽然每个个体在遗传上是独一无二的,但是每个人更像他(她)的亲属,而不像其他无亲缘的个体。正如下面将讨论的一样,多基因病的发病有家族性,因此,确定亲属间的遗传关系相当重要。图 4-5 的家系表示了这种关系。在符号内的数字代表了他与先证者的亲属级别,先证者用箭头和实心符号表示。先证者与一级亲属之间有 1/2 基因相同,与二级亲属有 1/4 基因相同,与三级亲属有 1/8 基因相同。值得注意的是,共享基因的分数与亲缘系数(r)相同(表 4-5)。配偶和其他通过婚姻建立的亲属可能是无亲缘关系,同时他们同先证者的遗传关系与群体中随机选择的个体同先证者的遗传关系无明显差别。

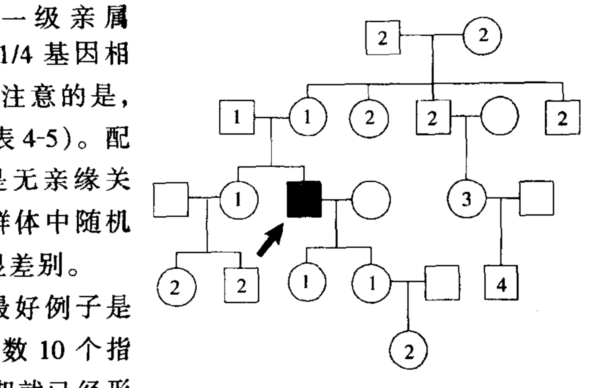


图 4-5 用于表示亲属级别的家系。在此图中,符号内的数字表示其与先证者(黑色符号)的亲属级别。

在人类,“纯粹”的多基因性状的最好例子是指尖嵴纹计数,即依据特定的规则,计数 10 个指尖总的嵴纹数。指尖嵴纹在发育的早期就已经形成,受环境因素影响很小。它的多基因特性表现为在群体中呈正态分布,亲属间指尖嵴纹计数的相关程度与他们共享的基因数非常近似(表 4-7)。

表 4-7 亲属间指尖嵴纹计数的相关系数的观察值与通过共享基因比例得到的期望值之间的比较*

关 系	相关系数观察值	相关系数期望值
单卵双生	0.95 ± 0.07	1.00
异卵双生	0.49 ± 0.08	0.50
同 胞	0.50 ± 0.04	0.50
父母 - 子女	0.48 ± 0.04	0.50
配 偶	0.05 ± 0.07	0.00

* 引自 Carter CO: Genetics of common disorders. Br Med Bull 25:52 ~ 57, 1969.

(一) 常见病和畸形的遗传因素证据

流行病学研究通过比较在遗传相关个体间和在一般群体内疾病的发病率,证明了遗传因素在常见病和畸形中的重要作用。首先,特定疾病或畸形的患者亲属发病率明显增高("家族倾向"),并且与亲属级别成正比。这一点通过对欧洲和北美洲地区伴有或不伴有腭裂的唇裂发病率的大规模家系研究得以解释(图 4-6)。在那里,该病的发病率约为 1/1 000(表 4-8)。亲属中唇裂的发病率明显高于一般人群,然而,受累亲属的比例既不符

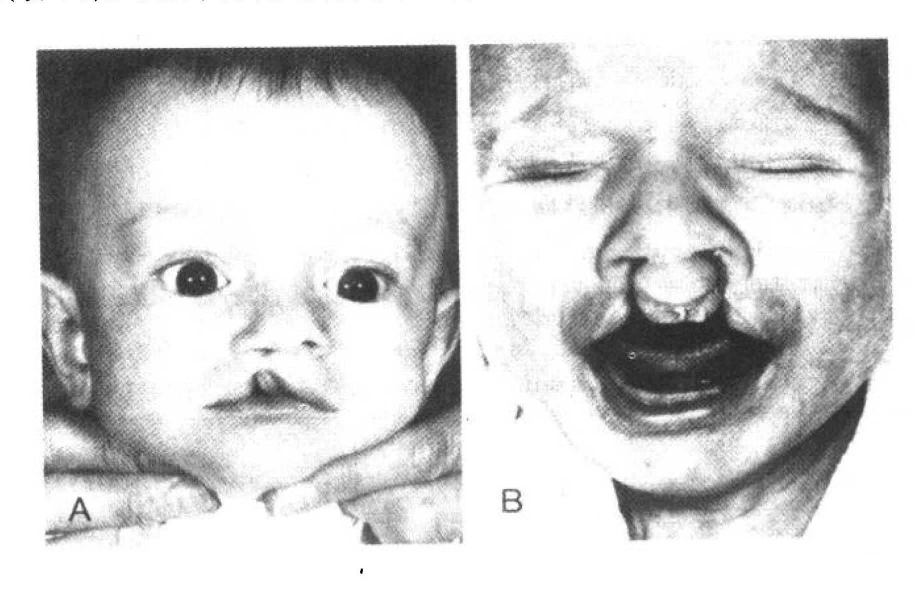


图 4-6 唇裂 ± 腭裂患儿。A:一侧唇裂患儿 B:双侧唇裂和腭裂患儿(引自 Ross RB, Johnson MC. Cleft lip and palate. Baltimore: Williams & Wilkins, 1972:131,41.)。

表 4-8 唇裂(±腭裂)发病率的家系研究*

亲 属	受累亲属百分数	同群体相比亲属的发病率
一级		
同胞	4.1	×40
子女	3.5	×35
二级		
姑和叔	0.7	×7
侄儿和侄女	0.8	×8
三级		
一级表兄妹	0.3	×3

合常染色体显性遗传,也不符合常染色体隐性遗传。二级亲属同一级亲属相比发病风险迅速降低,而三级亲属同二级亲属相比发病风险降低得更多。这与常染色体显性遗传中,亲属级别每降低一级,发病风险降低 1/2 的情况相比,降低的程度更大。受累先证者的同胞和子女的患病风险相同,这一点与常染色体隐性遗传不同。受累个体亲属患病风

险的增高用参数 λ_R 来表示，λ_R 是受累先证者亲属的患病风险除以一般群体的患病风险。下脚标 R 表示亲属的类型，如 λ_S 指的是同胞的患病风险。

然而，值得注意的是，不是所有家族性必定都是遗传的，家庭也面临共同的环境。父母之一为医生的医学生所占的百分比要高于从一般人群中医生数量得出的期望值。在 Cardiff 的威尔士大学的研究表明，综合考虑身体的、情感的、社会的和经济的因素，如果一个人的一级亲属中有医生，他成为医学生的“风险”提高 61 倍；若二级亲属中有医生，他成为医学生的“风险”提高 11 倍。然而，很难证明，“成为医学生”仅仅是由遗传因素决定的性状。

另一个更有说服力的证据，并支持遗传因素在常见疾病或畸形致病过程中作用的研究为双生子研究。如果性状(疾病或畸形)是完全遗传的，则 100% 的单卵双生子或遗传上一致的双生子(monozygotic, MZ)的性状应该是一致的(即都具有该性状或都不具有该性状)；异卵双生子或遗传上不一致的双生子(dizygotic, DZ)，由于他们共享 50% 的基因(但有完全相同的环境)，则 50% 的异卵双生子的性状应该是一致的。但是，如果性状是完全不遗传的，则单卵双生子和异卵双生子性状的一致程度大约相等，并明显小于 100%。最后，如果性状是多因素的并有明显的遗传因素时，则单卵双生子性状的一致程度小于 100%，但要明显高于异卵双生子的性状一致程度。几种有明显遗传因素的常见疾病和畸形的数据列于表 4-9。

表 4-9 在常见畸形和疾病中单卵双生子和异卵双生子的一致性程度

性 状	一致性	
	单卵(MZ)	异卵(DZ)
	%	
唇裂 ±腭裂	40	4
幽门狭窄	22	2
精神分裂症	46	14
胰岛素依赖性糖尿病	30	6

第三，一些罕见的、呈孟德尔方式遗传的先天缺陷或疾病说明，遗传因素能够引起特定的表型。例如，一种非胰岛素依赖性糖尿病——MODY(Maturity onset diabetes of the young)为常染色体显性遗传性状。部分胰岛素依赖性糖尿病为罕见的孟德尔疾病。

考虑常见的多因素疾病或者畸形的遗传基础时，必须牢记两点。第一，遗传决定的不是疾病本身，而是这种疾病的遗传易感性。有遗传易感性的个体能否发病，取决于各种遗传因素和环境因素的相互作用，包括饮食、活动、毒素等。第二，不同的遗传背景可以引起对相同疾病的易感性。所以，常见的多因素疾病的病因具有遗传异质性。例如，冠心病的危险因子包括高血压、糖尿病、高脂血症等，每个危险因子都有一套自己的遗传和环境的诱因。

(二) 连续分布的易患性中不连续的性状(疾病)：阈值模式

从上面的讨论可以看出，许多常见病和畸形的致病因素中包含了重要的遗传因素，

并且在疾病的形成过程中，环境诱因也同样存在。为了说明这些事实，以常见出生缺陷的流行病学研究为基础，对已知的疾病建立一个连续分布的遗传易患性模式。超过特定遗传易患性阈值的部分个体，尤其当存在适当的环境诱因时，将会发病。这种阈值模式说明了连续分布的遗传易患性中存在不连续的性状或疾病（图 4-7）。这种模式假定患者位于易患性分布的最右侧。因此，在同一条件下，与患者有 1/2 基因相同的一级亲属，其易患性的分布移至一

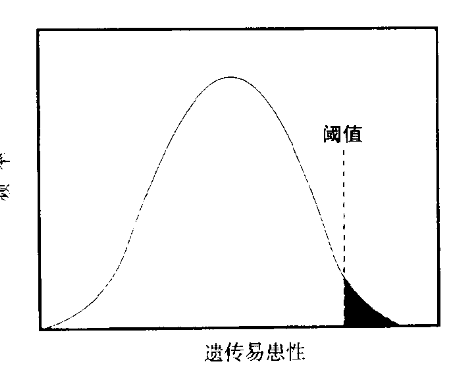


图 4-7　多因素遗传的阈值模式。

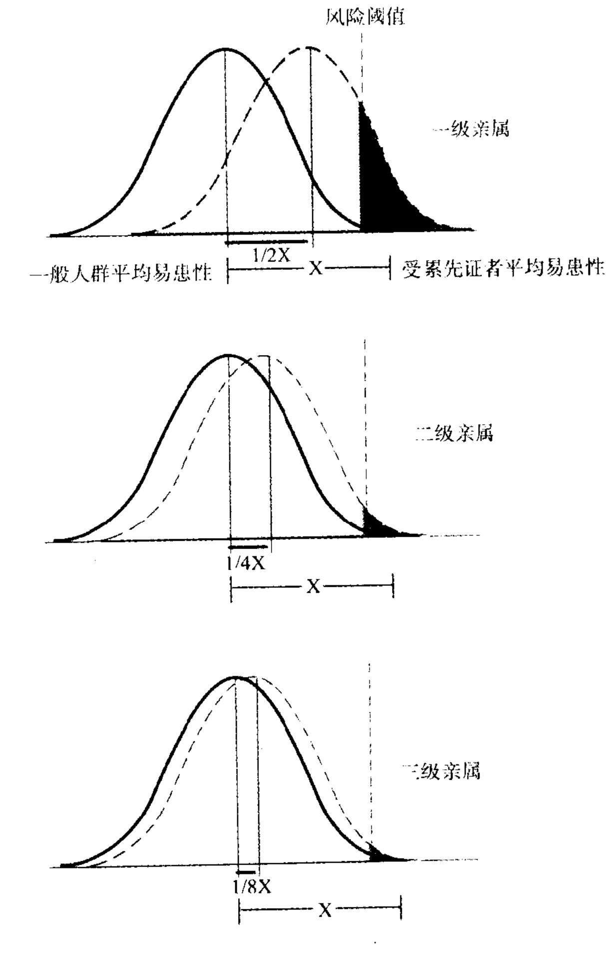


图 4-8　多因素阈值模式：亲属间遗传易患性的分布。先证者亲属遗传易患性的分布用虚线下灰色阴影区域表示；X 表示患者与一般人群易患性均值的差异（引自 Carter CO. Multifactorial genetic disease. Hosp Pract 1970;5:45～59.）。

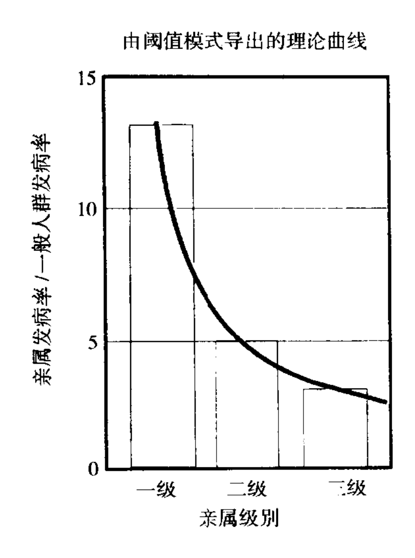


图 4-9　多因素阈值模式：预示亲属的再发风险（引自 Carter CO. Multifactorial genetic disease. Hosp Pract 1970; 5:45～59.）。

般人群易患性的右侧（图 4-8）。与患者有 1/4 基因相同的二级亲属（其余基因可认为来自一般人群），其易患性的分布靠近一般人群的易患性均值，但仍在其右侧。与患者仅有 1/8 基因相同的三级亲属，其遗传易患性的分布更靠近一般人群的易患性均值。在每种情况下，随着亲属级别的降低，一级、二级和三级亲属超过一定遗传性状阈值的百分比进行性减小。根据正态分布的几何特性可以推论一级亲属的患病风险明显增高，二级亲属的患病风险适度增高，而三级亲属的患病风险仅轻微增高（图 4-9）。

（三）多基因病或畸形再发风险的多因素阈值模式预测

多因素阈值模式能对患者亲属的疾病或畸形的再发风险进行多种预测，这些预测不同于孟德尔式遗传。

（1）*再发风险代表平均风险，在不同家庭中各不相同。*多因素阈值模式假定患者的易患性超过群体遗传易患性分布阈值，但它们的准确位置不能直接确定。平均再发风险预测是根据经验数据获得的，如表 4-8 所示，它不像孟德尔式性状那样由特定的遗传模式或方式决定。这个阈值模式还预示出：随着亲属级别的降低，亲属的再发风险迅速降低，如图 4-8 和图 4-9 所示。表 4-8 中唇裂 ±腭裂患者的一级、二级、三级亲属的再发风险及表 4-10 中几种常见的先天畸形的再发风险就符合这一点。

表 4-10　常见先天畸形的家族模式*

畸　形	群　体 发病率	同一般群体相比亲属的发病率			
		单卵双生子	一级亲属	二级亲属	三级亲属
唇裂 ±腭裂	0.001	×400	×40	×7	×3
畸形足	0.001	×300	×25	×5	×2
神经管缺陷	0.002		×8		×2
先天性臀错位 （只有女性发病）	0.002	×200	×25	×3	×2
先天性幽门狭窄	0.005	×80	×10	×5	×1.5

* 引自 Carter CO: Genetics of common disorders. Br Med Bull 25:52～57, 1969 and Smith DW, Aase JM: Polygenic inheritance of certain common malformations. J Pediatr 76:653～659, 1970.

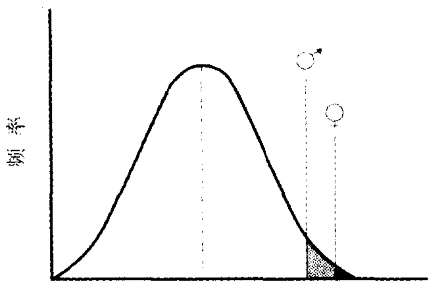
（2）*患病风险随受累亲属数目的增加而增高。*对于常染色体隐性疾病如囊性纤维化而言，无论这个家庭已有一个、两个或几个患儿，患者同胞再发风险仍为 25%。相比之

下，对于唇裂伴有或不伴有腭裂的患儿，其同胞的再发风险约为4%；如果两名一级亲属如两名同胞或双亲之一和一名同胞均为患儿时，其再发风险将大于10%。事实上，另一名患者的存在并不改变患病风险本身，但是，它提示这个家庭位于遗传易患性分布曲线右侧偏远处。因此，这个家庭的再发风险高于平均再发风险。

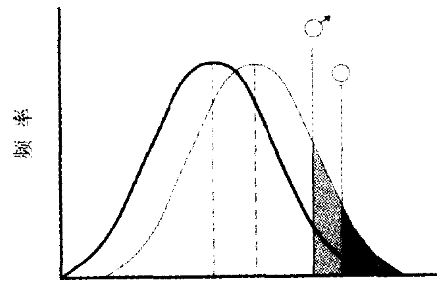
(3)*患病风险随着畸形或疾病的严重程度而增高*。单侧唇裂不伴有腭裂的患儿，其同胞的再发风险约为2.5%，而双侧唇裂伴有腭裂的患儿，其同胞的再发风险约为6%。这反映出缺陷越严重，潜在的遗传易患性越大，因此，严重患者位于分布曲线右侧偏远处。

(4)*随着一般群体中疾病或畸形发病率的降低，先证者亲属患病风险的差异增加*。畸形在一般群体中的发病率越高，则遗传阈值越低，导致整个群体和受累个体平均遗传易患性的差异减小。因此，患者亲属与一般群体间平均遗传易患性的差异较小。与先证者亲缘关系较近的亲属，其患病风险的增高程度相对减少。

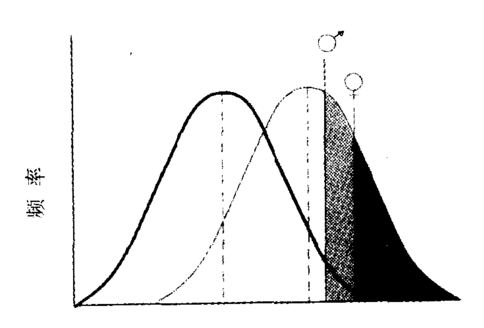
(5)*当受累先证者的性别比例显著失调时，发病率低的性别，其后代发病的风险相对较高*。先天性幽门狭窄是由于肌肉增生造成的胃出口阻滞，男孩发病率比女孩发病率高5倍。这意味着女孩遗传易患性的阈值高于男孩(图4-10)。因此，受累女孩的遗传易患性较高，她的亲属比受累男孩的亲属具有较高的再发风险。事实上，受累女性的子女患有这种畸形的可能性是受累男性子女的3倍。但是，后代中男性发病风险的绝对值仍高于女性(表4-11)。



一般人群的遗传易患性



男性先证者的一级亲属



女性先证者的一级亲属

图4-10 多因素阈值模式：先证者及其亲属幽门狭窄发病率的性别差异。男性和女性发病风险阈值用实心垂直线表示。男性和女性先证者亲属遗传易患性分布曲线用细曲线下的灰色浅阴影区表示，患者用灰色深阴影区表示。详见正文(引自 Thompson M. Genetics in medicine. 4th ed. Philadelphia: WB Saunders, 1986:217.)。

尽管上面的研究无法证明阈值模式的合理性，而实际上其他多因素疾病的非阈值模式已经建立，但这种多因素阈值模式为常见病和先天畸形的家庭遗传咨询提供了有用的框架。除了已知的孟德尔式遗传外，这些患者经常得到错误的咨询，要么(当他们没有较高的再发风险时)告诉他们有较高的再发风险，要么(当他们确实有较高的再发风险时)告诉他们根本没有再发风险。对于大多数多因素疾病如胰岛素依赖性糖尿病、唇裂、先天性心脏病而言，先证者的子女或者同胞的再发风险约为5%或更小。

表 4-11 子女患有幽门狭窄的比例*

先证者	子女	
	儿子	女儿
	%	
父	5.5	2.4
母	19.4	7.3
群体发病率	0.5	0.1

* 引自 Carter CO:Genetics of common disorders. Br Med Bull 25:52~57,1969

(四) 常见复杂疾病遗传因素的鉴定

利用分布于整个人类基因组的高度多态标记的特征,寻找常见多因素疾病易感基因已成为可能(详见第五章和第九章)。两种主要方法是**相关分析**(association analysis)和**连锁分析**(linkage analysis)。前者主要寻找群体中某一遗传位点的特定等位基因与疾病的相关程度。通常这种方法需要利用**候选基因**(candidate gene),即编码已知蛋白或认为与疾病形成有关的蛋白的一类基因。发现特定的等位基因在患者中出现的频率明显高于正常个体,提示该等位基因产物在疾病的致病过程中发挥作用。例如,*apoE* 基因的 ε4 等位基因可能是迟发型老年性痴呆病的一个重要的危险因子。相比之下,连锁分析是寻找**家系**(family)内某一标记基因座与疾病的共传递情况,而不考虑该基因座的等位基因。连锁分析直接证明一个遗传基因座或与之相邻近的遗传基因座在疾病形成过程中的作用。最后,建立人类多因素疾病的动物模型可以发现疾病的遗传因素,多用于人类同源序列的研究。

胰岛素依赖性糖尿病

Ⅰ型糖尿病(胰岛素依赖性糖尿病,insulin-dependent diabetes, IDDM)已成为多因素疾病遗传分析的一个范例。IDDM 以前称为青少年期糖尿病,群体发病率为 0.4%,高发年龄在 12 岁。它的主要特征为胰岛 β 细胞的自体免疫性破坏,导致胰岛素的绝对不足,而需要外源性胰岛素来维持生存。IDDM 是心脏病、肾功衰竭、失明的重要危险因子。IDDM 的发病具有明显的家族性,10%的患者有一个患病同胞;糖尿病患者同胞和 IDDM 男性的子女患病风险高于群体患病风险 10 倍以上;IDDM 单卵双生子的一致性约为 25%~30%,大约比异卵双生子的一致性高 5 倍。

研究胰岛素依赖性糖尿病的遗传因素最主要的方法是分析家系中受累同胞对,主要观察某一染色体基因座特定的等位基因的传递一致性(identical-by-descent, IBD)频率,即是否从一个共同祖先如双亲遗传下来,然后在一个遗传基因座的 IBD 值可以与随机期望值相比较。在随机分离的情况下,两个同胞在任何基因座的 IBD 值可能为 0,1,2,概率分别为 25%,50%和 25%(图 4-11)。

相关性研究和连锁分析均鉴定了 HLA 区域一个称为 *IDDM*-1 的基因座对决定 IDDM

易感性起主要作用。编码Ⅱ类分子 HLA-DQ 和 HLA-DR 的基因尤其重要，这两类分子与多肽抗原递呈给 T 淋巴细胞有关，从而在免疫应答中发挥关键作用。HLA DR3 和 DR4 杂合个体发病风险特别高。另一个重要的基因座称为 *IDDM*-2，是胰岛素基因本身(insulin gene itself, INS)；易感等位基因与胰岛素基因表达增高相关。利用基因组扫描技术(详见第九章)，目前至少已经发现了另外 11 个遗传基因座与 IDDM 易感性有关。通过分析非肥胖性糖尿病(non-obese diabetic, NOD)小鼠，已经发现至少 10 个基因座与小鼠的 IDDM 易感性有关，但至今仍未在人类鉴定与之相对应的新的易感基因。

通过分析 *HLA* 和 *INS* 基因座的多态性，可以在无 IDDM 家族史的群体中确定糖尿病的高危人群。而在糖尿病家系中，该分析可以推断出子女有 25% 的发病风险。最后，通过遗传标记研究结合胰岛自身抗体及 β 细胞功能检测，我们可以鉴定出糖尿病家系中子女 IDDM 非常高的发病风险。

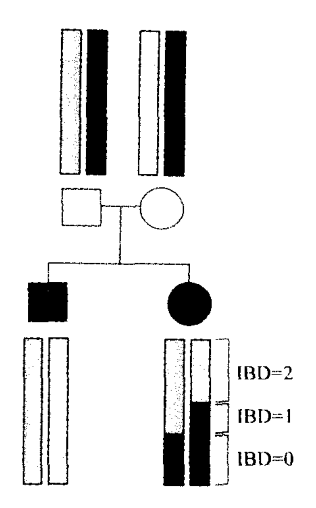


图 4-11　受累同胞共享等位基因的传递一致性。双亲特定区域的染色体可以用高度多态的遗传标记来表示，这将有助于判断双亲的同源染色体和后代从亲本遗传得到的片段。比较后代的染色体可以判断同胞间的 IBD。同胞间可以共享亲本的 2 个拷贝(2)，1 个拷贝(1)或不共享拷贝(0)。通过测定若干家系中每个遗传位点的 IBD 程度，可以找到高于期望值的区域，该区域可能有易感基因毗邻(引自 Courtesy of Professor Aravinda Chakravarti, Department of Genetics, Case Western Reserve University.)。

(五) 多因素遗传的意义

多因素阈值模式提示疾病的环境诱因很可能对遗传易感个体有重大影响。因此，重点研究具有高度遗传风险的个体时，寻找该疾病的环境诱因应是最有成效的。在多因素疾病中，这些个体的发现将有助于发现多因素疾病的环境因素。相应地，通过改变这样个体的生存环境而进行的医疗干预可以集中于具有高度遗传风险的个体，使他们从医疗干预中受益。也许长期性更重要，将有可能鉴定常见病易患性的遗传标记进而探寻这些疾病的致病机制。这又将导致新的治疗方法的诞生。

尽管多因素疾病遗传学的研究经常假定有多个基因，每个基因起微小作用，但一些非孟德尔式遗传病如 IDDM 很有可能是相对较少的几个基因作用的结果，其中一些基因对表型的影响相当大。因此，这些疾病可能是寡基因病而不是多基因病的。如上所述，这一点对于进行遗传咨询和群体筛查、寻找易感个体有重要意义。

推　荐　读　物

Bell JI. Polygenic disease. Curr Opin Genet Devel 1993;3:466 ~ 469.

Carter CO. Genetics of common disorders. Br Med Bull 1969;25: 52 ~ 57.

Carter CO. Multifactorial genetic disease. Hosp Pract 1970;5:45 ~ 59.

Cavalli-Sforza LL, Bodmer WF. The genetics of human populations. San Francisco: WH Freeman, 1971.

Emery AEH. Methodology in medical genetics. An introduction to statistical methods. 2nd ed. Edinburgh: Churchill Livingstone, 1986.

Harris H. The principles of human biochemical genetics. 3rd ed. Amsterdam: Elsevier/North-Holland, 1980.

Keats B. Population genetics. In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics. 3rd ed. New York: Churchill Livingstone, 1996:347 ~ 357.

King RA, Rotter JI, Motulsky AG, eds. The genetic basis of common disease. New York: Oxford University Press, 1992.

Lander ES, Schork NJ. Genetic dissection of complex traits. Science 1994;265:2037 ~ 2048.

Lathrop GM, Terwilliger JD, Weeks DE. Multifactorial inheritance and genctic analysis of multifactorial disease. In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics. 3rd ed. New York: Churchill Livingstone, 1996: 333 ~ 346.

Risch NJ. Genetic epidemiology. In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics. 3rd ed. New York: Churchill Livingstone, 1996:371 ~ 382.

Todd JA. Genetic analysis of type Ⅰ diabetes using whole genome approaches. Proc Natl Acad Sci USA 1995;92: 8560 ~ 8565.

Vogel F, Motulsky AG. Human genetics. 3rd ed. Berlin: Springer-Verlag, 1996.

第五章　分子遗传学:基因结构、调节和操作

正如我们现在的医学理论和实践依赖于精通人类解剖学、生理学以及生物化学的知识那样，将来处理疾病也需要对人类基因组的分子解剖学、生理学和生物化学的详细了解。我们将必须有熟悉染色体和基因分子解剖学的内科医生，就像心脏外科医生熟悉心脏和循环系统的结构和功能一样。

——Paul Berg, M.D. 诺贝尔奖演说, 1981

第二章概述了遗传信息从 DNA 至 RNA 最后到蛋白质的传递。若要了解人类遗传疾病的真正本质，就需要更精确地掌握人类基因的结构和功能。这就需要我们进一步深入研究遗传分析技术，统称或笼统地称为"重组 DNA 技术"。

一、基因的一般结构和基因家族

人类 23 对核染色体上分散存在着大约 5 到 10 万个基因。这些基因最小的不超过几百个碱基对的 DNA;而最大的已知人类基因(抗肌萎缩蛋白)长度为二百多万个碱基对。图 5-1 显示一个模式基因结构，模拟成年人血红蛋白基因，对此我们将在第六章详细说明。血红蛋白是一个由两个 α 蛋白链和两个 β 蛋白链组成的四聚体。编码这些蛋白的基因分布在不同的染色体上:α 基因在第 16 号染色体上,β 基因在第 11 号染色体上(图 5-2)。α 基因长约 800bp,β 基因长度大约是它的二倍。然而它们的蛋白产物却相似,α 长为 141 个氨基酸;β 为 146。

大多数真核细胞的一个显著特征是存在**内含子**(intron)。内含子是位于基因内的 DNA 片段，它能够转录成 RNA，在翻译成蛋白质之前被剪接掉。图 5-1 中空心部分就是内含子。该基因示意图的黑色实心部分是含有编码信息的区域。在该基因的两端也有非翻译区，在 5-1 图中用黑色条纹状框表示。5′非翻译区是被 RNA 聚合酶转录成信使 RNA、而不被翻译的一段序列。一般情况 RNA 上的第一个 AUG 为翻译开始点，AUG 被称为起始密码子，位于它 5′端的序列不编码蛋白质。3′端非翻译区是在 mRNA 上终止密码子(UAG, UGA, 或 UAA, 见表 2-2)之后的一段序列，因此也不翻译为蛋白质。编码序列、5′和 3′非翻译区共同构成了**外显子**(exon)，剪接后连在一起形成成熟的信使 RNA，即 mRNA。在血红蛋白 α 基因中，外显子 1 编码 5′非翻译区和 1～31 氨基酸;外显子 2 编码第 32～99 氨基酸;外显子 3 编码第 100～141 氨基酸以及 3′非翻译区。一个基因内的内含子数量变化很大，偶尔有些基因内只有一个外显子;而有些基因内含子可多达 100 个。

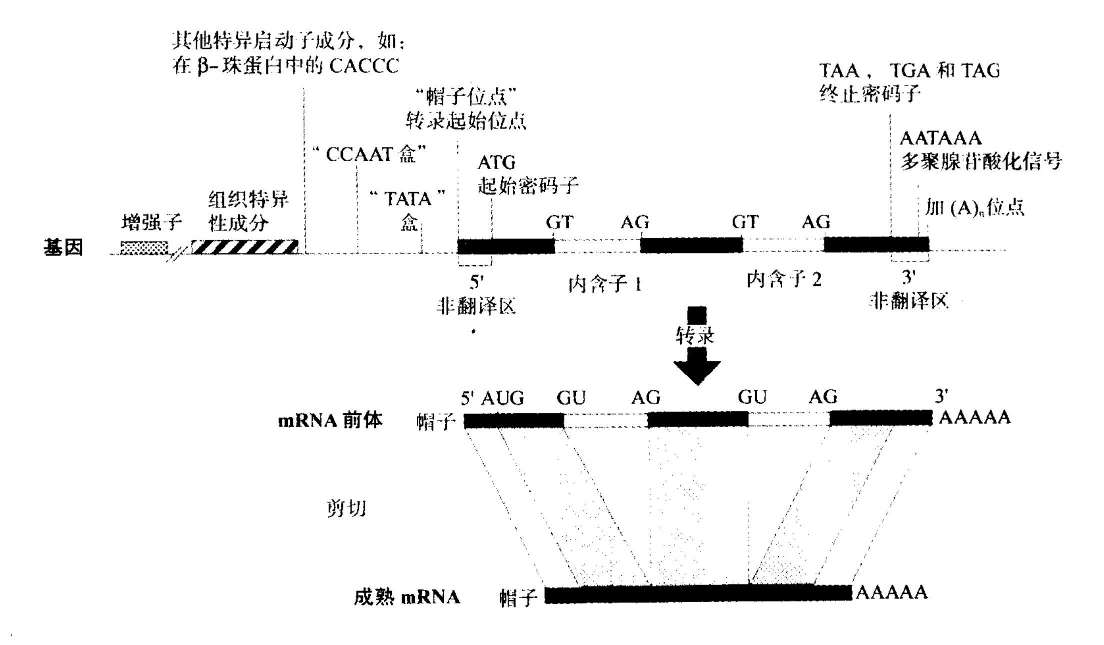


图 5-1 一个理想化基因的结构及表达，包括各种启动子元件，一个增强子和该基因的转录区。该例以血红蛋白基因为模型。其中的两个内含子均以 GT 起始，以 AG 终止。在形成成熟 mRNA 时内含子被剪接。如文中所述，不仅启动子成分而且增强子都能具有组织特异性。

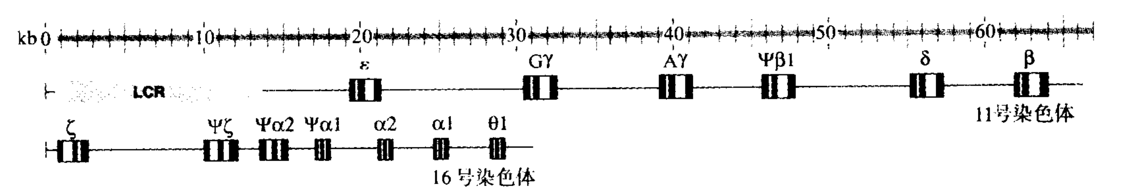


图 5-2 11 号染色体短臂 β 珠蛋白基因簇图和 16 号染色体短臂 α 珠蛋白基因簇图。每一例中各基因按照它们在发育过程中激活的序列从 5′端向 3′端排列。基因 ζ 和基因 ε 是胚胎基因；Ψζ、Ψα2 和 Ψα1 以及 Ψβ1 是假基因，它们由于获得了一些导致失活的突变而不能产生功能性的珠蛋白链。θ 是一个珠蛋白样基因，没有发生导致失活的突变，它的功能未知。作为胎儿期主要的血红蛋白，血红蛋白 F 包含 α 链和 γ 链；两个 α 基因(α1 和 α2)编码相同的蛋白；两个 γ 基因(Gγ 和 Aγ)在第 136 密码子上其产物相差一个氨基酸。成年期主要的血红蛋白是 α2 β2，也有少量的 α2 γ2。基因内白框表示内含子。请注意在各基因之间有许多"间隔"DNA，其中一些含有重复序列。LCR 基因座控制区将在第六章讨论。

有时功能相似的基因成簇存在。图 5-2 显示人类的 α 和 β 珠蛋白基因簇。在 β 珠蛋白基因附近有 4 个有其他功能的基因，分别是 ε、Gγ、Aγ 和 δ，它们也编码血红蛋白，但是它们在发育的不同阶段表达。与 β 珠蛋白基因相似，α 珠蛋白基因也是基因簇的一部分(实际上有 2 个基因：α1 和 α2，编码序列相同)。ζ 基因是一个胚胎基因；标记 θ1 的基因其功能未知。在 β 珠蛋白基因簇中，另有一个标记 Ψβ1 的基因；在 α 珠蛋白簇中分别也有标记 Ψζ、Ψα1 和 Ψα2 的基因。这些所谓**假基因**(pseudogene)，标记为 Ψ，它们有表达基因的部分结构，被认为曾经有功能，但在进化过程中获得了一个或多个突变而丧失了产生蛋白产物的能力。

在图 5-2 中与珠蛋白基因簇一样明显的是在转录的基因间常有相当长的 DNA 片段。在 β 珠蛋白附近分析出 80kb(1kb = 1 000bp)长的 DNA 中，只有 12%实际上能被转录，仅 2.5%编码蛋白。非转录 DNA 被称为“**基因间的**”(intergenic) DNA。靠近表达基因的一些基因间 DNA 序列对于基因表达调控至关重要(后面我们将会看到)，但绝大部分基因间 DNA 似乎不重要，没有已知功能。与此相似，很多基因含有很大的内含子，绝大部分看来也是不重要的。例如，抗肌萎缩蛋白基因的 2 400kb DNA 中内含子占 99%以上。在基因间的 DNA 中有时也在内含子中有大量的重复序列，它们数以千计的拷贝广泛散布在基因组中，没有明显已知功能。其中最常见的所谓 *Alu* 的重复序列，大约长 300 bp，在人类基因组中出现了 50 万次。因为它们数百万年前就散布在基因组中，*Alu* 序列已发生了变化，所以一个 *Alu* 重复序列与另一个大约有 80%相同。通常一个大于 30 kb 的 DNA 片段内至少有一个 *Alu* 序列。

二、基因表达调控

除极少数显著的例外，人体所有细胞都包含完整的基因组。但是在任一个特定组织中只有少部分基因表达。因此，基因表达的调控对于了解人类生物学几乎所有方面都至关重要。

一般来说，基因的成熟蛋白产物发挥基因的功能。成熟蛋白表达水平受如下因素影响：(a)基因转录成 RNA 的速率；(b)RNA 的加工；(c)mRNA 从细胞核向细胞质转运；(d)在细胞核糖体中 mRNA 翻译成蛋白质的速率；(e)mRNA 降解速率；(f)蛋白翻译后修饰；(g)蛋白降解速率。所有这些调控机制在特定场合都能涉及到。然而或许最经济的，也是在真核细胞中广泛应用的方法就是在其最早水平控制产生蛋白，也就是基因转录水平。图 5-1 是一个理想化人类基因控制元件示意图。这些重要的序列元件是通过各种方法鉴定的，包括：突变分析，进化比较，应用将基因转移至培养细胞或转基因小鼠进行的功能分析等等。

(一) 启动子

启动子较笼统地确定为紧靠基因 5′端的 DNA 序列元件，它们与 RNA 聚合酶以及转录系统其他成分相互作用。启动子决定转录起始位点以及控制 mRNA 的数量甚至有时决定转录的组织特异性。一些重要的启动子元件一般位于基因 5′端 100 ~ 200bp 范围内，但有些启动子可长达数千个 bp。

很多人类基因在转录起始点 5′端上游 25 ~ 30bp 处有一个保守的“**TATA 盒**”序列，可能参与转录起始位点的准确定位。其上游距起始点 75 ~ 80bp 的区域，常有一个“**CCAAT 盒**”序列，与“TATA 盒”相比较为少见。在那些有“CCAAT 盒”的基因中，至少在基因转移实验中证实，它的存在看来是定量高效转录所必需的。值得注意的是，编码几乎所有细胞都有编码酶的某些“持家”基因，却常常缺失这两个“盒”，并有富含 C 和 G 核苷酸的启动子。无“TATA 盒”基因的转录起始点在 10 ~ 20bp 范围内常显异质性。在不断延长的

mRNA 链 5′端加上一个经特殊修饰的核苷酸,7 - 甲基鸟苷,称为帽子。所以转录起始点也经常被称为“**帽位点**”(cap site)。

(二) 剪接

如前所述,大多数真核细胞基因具有由内含子分隔的编码区,必须经过称为剪接的过程去掉内含子,产生成熟的 mRNA 才能翻译为有功能的蛋白质。尽管内含子的功能仍然不清,但剪接的机制正在被认识。人们在内含子的起始和末端发现了一些核苷酸序列(图 5-1)。内含子几乎总是始于 GT(剪接供体),终于 AG(剪接受体),其他相邻的碱基往往形成一定的序列,称为一致序列(图 5-3)。这些一致序列虽然必要,但不完全能满足剪接体识别之用。人们会发现在一些转录的基因中一致剪接供体和受体序列并未用上。有趣的是,突变使正常剪接信号失活有时能激活这些“隐蔽”的剪接信号。

5′外显子　　内含子　　3′外显子

$\frac{C}{A}AG\ GT\frac{A}{G}AGT \text{—//—} \binom{T}{C}_{\geq 11} N\frac{C}{T}AG\ G$

图 5-3　5′和 3′剪接位点的一致序列。在内含子开头的 GT 和结尾的 AG 几乎是不变的,而在一致序列的其余部分经常能见到变异。标有 N 的部位可以是任何一种核苷酸(A、C、G 或 T)。

特殊剪接供体“发现”相应的剪接受体的机制仍然不清。从 5′到 3′端的扫描模式是一种可能,但它与具有某种剪接受体突变的剪接模式不一致。然而,考虑到某些基因如胶原基因(见第七章)虽有多达 50 个分离的内含子,但供体和受体仍然总能正确地连接,故随机搜索模式也不合理。

(三) 多聚腺苷酸化

大多数编码蛋白的信使 RNA 都具有在它们的 3′端附有一串大约 200 个腺苷酸残基的长链(多聚腺苷酸化)的特征。这类 mRNA 一致的特点是在 3′非翻译区都有一个六核苷酸信号 AAUAAA,尽管在其附近其他序列对正确的多聚腺苷酸化也起作用。在 AAUAAA 信号下游 18 ~ 20bp 的部位加上 A 残基。“多聚 A 尾”似乎对 mRNA 从核中向外转运以及调节 mRNA 稳定性起一定作用。

(四) 增强子

增强子是具有以下特征的 DNA 序列:(a)它们增加附近一个基因的转录;(b)它们的转录增强作用能跨越相当长的距离并且相对地不受距离变化的影响;(c)即使序列颠倒仍然起作用。最早被定性的增强子源于 DNA 病毒如 SV40,它有一个符合上述标准的 72bp 二倍重复序列,并能在几乎所有被检的组织中增强许多基因的转录。最近,发现了

组织特异性增强子。例如位于免疫球蛋白基因内的增强子,它只在 B 细胞(能合成免疫球蛋白)中而不在其他组织中发挥作用。

(五) 转录因子

过去几年中,人们在转录水平了解基因表达的内在调控取得很大进展。包含 RNA 多聚酶和许多相关分子的转录起始复合物与启动子结合来进行 RNA 的合成。很多转录因子控制转录过程。转录因子是一类蛋白,它们与目标基因启动子或增强子中特定序列相结合来增强、降低或以其他方式调节基因的表达水平。现已开发出很有效的技术来研究转录因子如何结合 DNA 以及描述它们的功能。这些蛋白常含有直接结合 DNA 的特异结构域以及与其他蛋白相互作用的有关节段。

众多大家族的转录因子与启动子和增强子中的特定 DNA 靶序列相互作用;转录因子之间也相互作用。这些作用决定了所有高等生物必需的复杂的发育和组织特异性基因表达。最近,人们发现一些转录因子的遗传性异常在一些人类疾病中起重要作用。

三、操作的工具:重组 DNA 技术

对重组 DNA 技术的总体认识对研究现代医学遗传学十分必要。当这项技术刚刚呈现在我们面前时很难正确评价它的威力。然而纵观本书的后半部分,人们就会了解这些技术为揭示人类遗传病的本质起着重要作用。还应该指出人们对于人类遗传病的革命性的认识源于近些年来重组 DNA 技术的发展。在现代医学遗传学几乎任一实验中起主要作用的聚合酶链式反应(PCR)技术,大约仅在 10 年前才发现。更近期的进展是应用定位克隆鉴定疾病基因以及在实验动物中破坏特定基因等方法(图 5-4)。

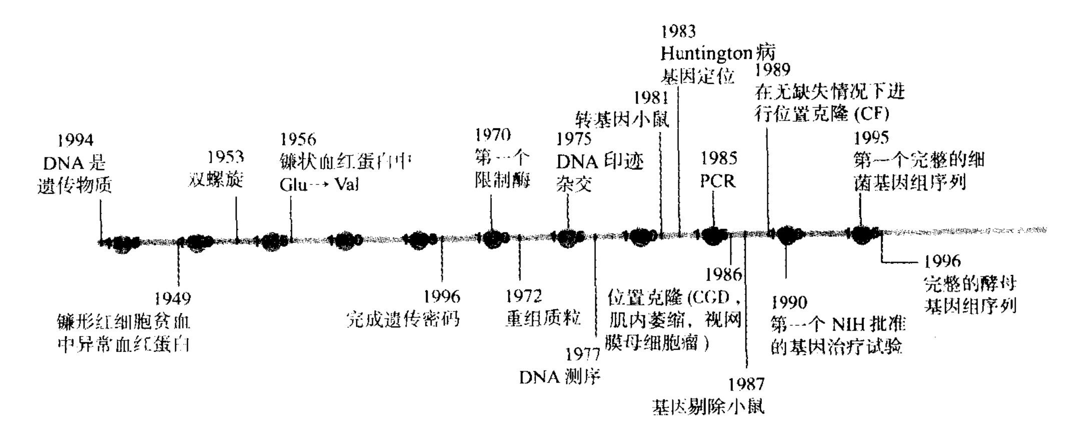


图 5-4 在现代分子遗传学发展中精选的突出事件。该时间线始于 1944 年 Avery、McLeod 和 McCarty 发现生物体内关键的遗传物质由核酸构成。请注意本章讨论的许多 DNA 重组技术标准方法都是最近一些年来的成果。CF:囊性纤维化;CGD:慢性肉芽肿病。

(一) 限制酶

限制酶也称限制性内切核酸酶,因为它们能在链内切断 DNA。它们是从细菌中纯化的一类酶,能在特殊的核苷酸序列上切断双链 DNA。如图 5-5 显示的例子。*Eco*R Ⅰ 的名称源于它是从大肠杆菌 RY13 菌株中分离出来的。它将切断 DNA 中任何带有 GAATTC 序列的部位。如图 5-5 所示,切断后留有一个突出的单链末端。这些突出的或"黏性"末端通常很有用,因为它们促使两个被同一限制性酶切断的 DNA 片段重新连接,连接由另一种称为 DNA 连接酶来完成。

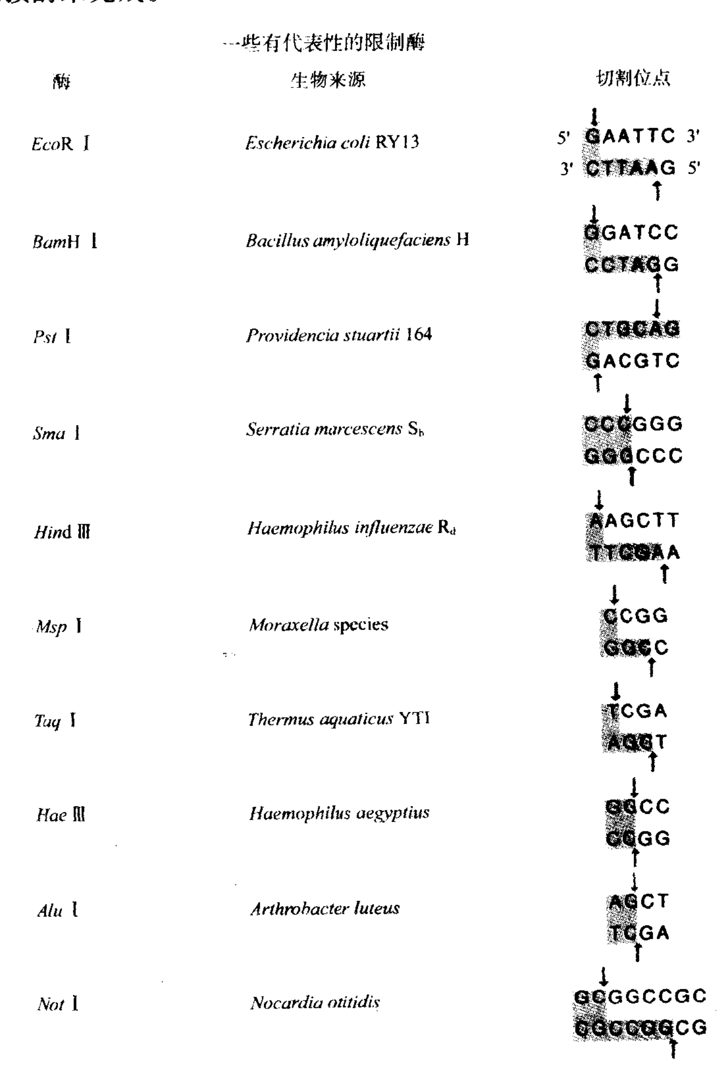


图 5-5 一些有代表性的限制酶、分离这些酶的细菌菌株以及它们识别的 DNA 序列。图上断裂位点的箭头表示实际切口位置。其中一些酶如 *Eco*R Ⅰ 切开后产生突出的黏性末端;而其他一些酶如 *Sma* Ⅰ 则产生平末端。

不是所有的限制酶都产生黏性末端。例如,图 5-5 中的 *Sma* Ⅰ 酶产生一个平末端。一些限制酶识别许多 4bp 序列(如 *Alu* Ⅰ 和 *Hae* Ⅲ),而 *Not* Ⅰ 实际只识别一个 8bp 序列,因此它在基因组 DNA 中切割相当少。一个特定限制酶的识别序列(限制性位点)在一段特异 DNA 中出现的频率部分地取决于它的长度。例如,由 *Eco*R Ⅰ 识别的 GAATTC 这样一个特殊的 6bp 限制性位点,在一个 DNA 随机片段中可能出现的频率约为每 4^6(或 4 096)个核苷酸一次,因为在这 6 核苷酸位点中,每一个位点都有 4 种可能性(A、G、C 和 T)。因此当 DNA 被 *Eco*R Ⅰ 酶切时,我们预计也许在平均 4 069bp 长度中产生片段。由于这些位点是随机分布,片段的长短将变化更大,可以远远小于也可明显大于 4 096 碱基对。与此相似,识别 4bp 序列位点的限制酶,如 *Alu* Ⅰ,将会平均以每 4^4 或 256bp 的间隔切断 DNA;识别 8bp 序列位点的限制酶,如 *Not* Ⅰ 将会平均每隔 4^8 或约 65 000bp 切断 DNA。实际上,因为在基因组中某些序列过多或过少地出现,一个特定限制酶位点的出现频率可能多于或少于预计的频率。

(二) 制备重组体

限制性酶能切割一个特定的 DNA 片段,但要大量制备这种片段,就必须把它插入一种"**载体**"(vector),这种载体就是能自主复制的 DNA 片段。最常用的载体是细菌**质粒**(plasmid)(带有抗生素抗药性基因的环状 DNA 分子)和**噬菌体**(bacteriophage,细菌病毒)。把一个特定的 DNA 片段插入这些载体之一,能制备毫克量的 DNA 用以详细分析基因结构。图 5-6 描绘出限制性酶和它们所产生的黏性末端怎样用来产生重组体。现已开发了许多种其他载体用于在其他类型细胞包括酵母和哺乳动物细胞中扩增克隆的 DNA。**反转录病毒**(retrovirus)经常用于介导外源基因进入哺乳动物细胞(见第十三章)。已开发出另一种被称为**酵母人工染色体**(yeast artificial chromosome,YAC)的载体特别用于携带很长的 DNA 片段(从数百 kb 到超过 1 Mb),实践证明这对基因制图非常有用(见第九章)。

(三) 互补 DNA(cDNA)

有时人们希望从基因转录的 RNA 而不是基因组 DNA 序列本身来着手研究。若利用重组 DNA 策略,单链 RNA 首先必须转变为双链的 DNA。这种转换由**反转录酶**(reverse transcriptase)来完成。图 5-7 描绘从单链 mRNA 到双链 DNA(所谓互补 DNA 或 cDNA)的转换。

(四) 分子杂交

在含有大量不同序列的混合物中检出所需的特异 DNA 或 RNA 序列常以分子杂交方法为基础。简单地说,杂交就是一个单链 DNA 或 RNA 通过碱基配对与它的互补单链结合,而不与无关单链序列结合的能力。图 5-8 表明这一过程。DNA 探针常用诸如放射性同位素等来标记,以便跟踪它们的存在。例如将许多独立的克隆固定在硝酸纤维素滤膜

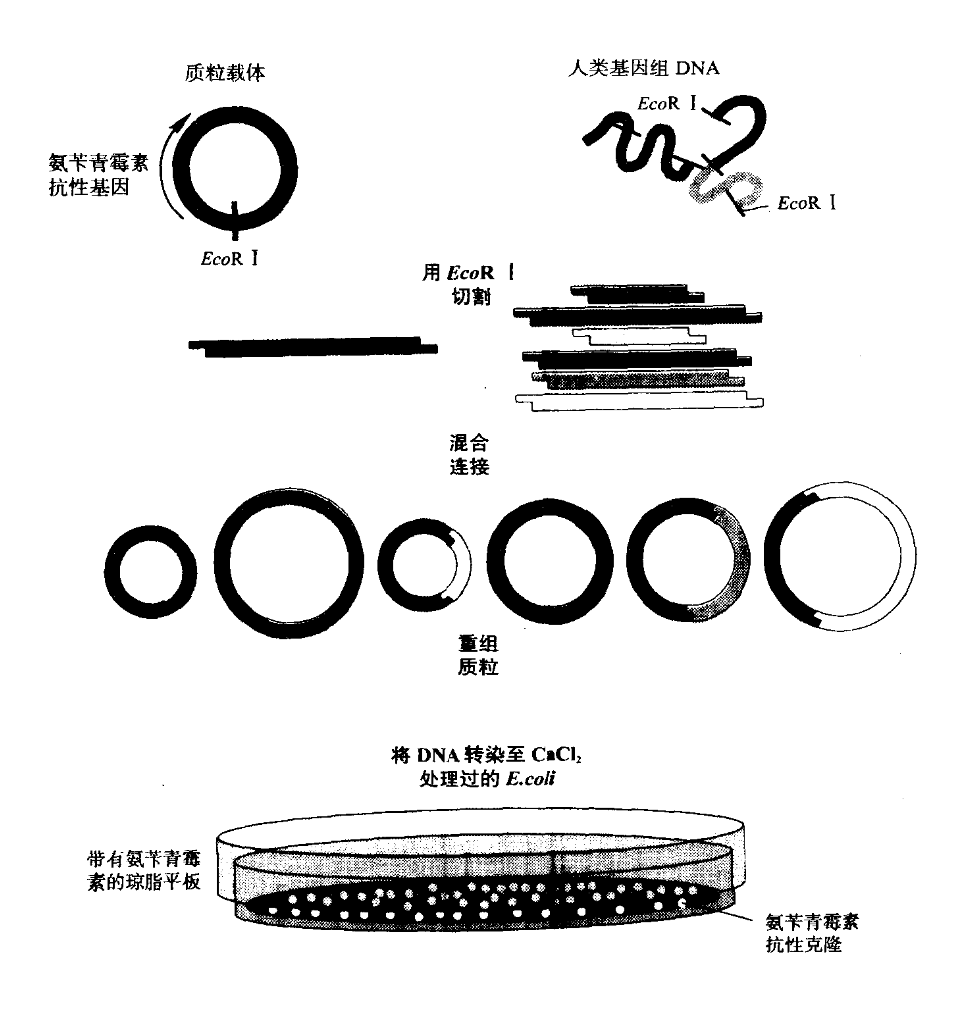


图 5-6　制备特异重组 DNA 分子的过程图解。在该例中，用限制酶 *Eco*R Ⅰ 分别切割含有目的片段的基因组 DNA 和带有氨苄青霉素抗性基因的细菌质粒载体。黏性末端被连接起来以后就产生了一系列重组分子，在这些分子中不同的基因组片段已被插入载体。再将该混合物转染氨苄青霉素敏感的大肠杆菌。在转染之前用氯化钙处理大肠杆菌以增加其对 DNA 的摄取能力。当它们覆于氨苄青霉素时，只有摄取了一个质粒的细菌才能形成菌落。每个氨苄青霉素抗性菌落是一个源于单个重组质粒分子的质粒克隆。有许多不同的方法能用来从培养皿中大批克隆或克隆文库中辨认和分离出所需特异重组克隆。这些方法包括图 5-9 中所示的用特异 DNA 探针来杂交筛选。

上，用标记的探针杂交，而后用滤膜使 X 线照片曝光，将会检出带有互补 DNA 序列的特定克隆。后文我们将会看到利用杂交鉴定特定核酸序列也就是决定 RNA 印迹法和 DNA 印迹法原理的基础。

（五）克隆基因

如图 5-9 所描绘，利用杂交克隆 cDNA 需要一种能在含有许多不同 DNA 序列的混合

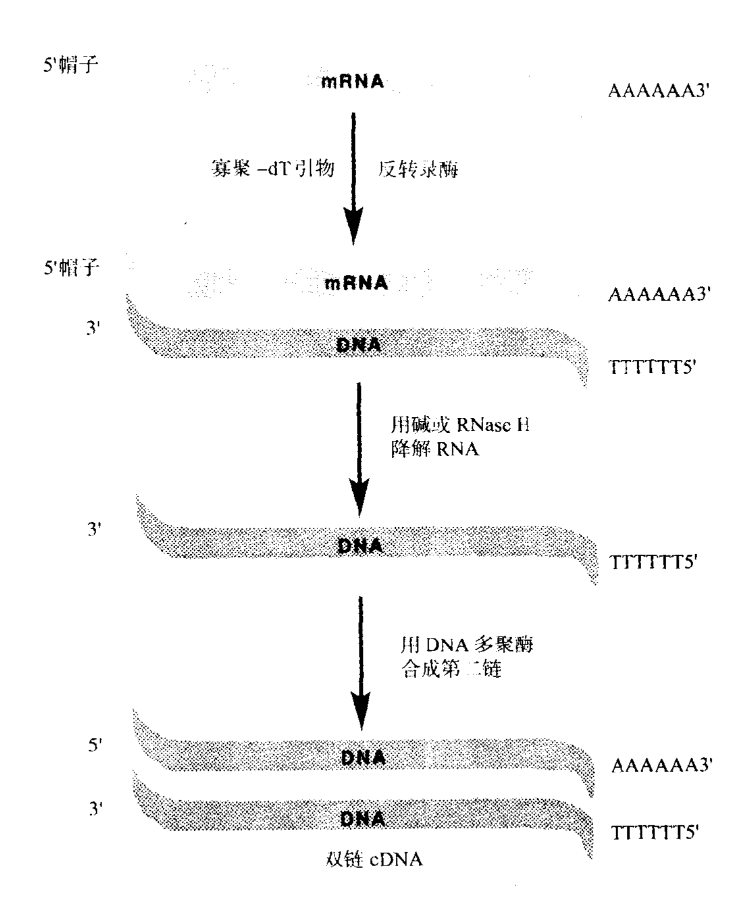


图 5-7 从单链 mRNA 制备双链 cDNA 的有关步骤。RNase H 是一种将在 RNA-DNA 杂合双链中切断并降解 RNA 链的酶。

物中识别所需克隆的试剂——**"探针"**(probe)。例如克隆突变的β珠蛋白基因并不难,因为我们可以利用正常的β珠蛋白基因序列。然而如果这个基因未被克隆,设计探针就是一个挑战性任务。如果该基因在一个特定组织中表达量非常高,从该组织中制备 cDNA 克隆就可能产生一些有正确插入物的克隆。例如α和β珠蛋白是红细胞的主要蛋白成分;以前成功地用网织红细胞(仍含有 mRNA 的早期红细胞)获得首批血红蛋白 cDNA。如果没有这样丰富表达的组织,就必须了解有关该基因的一些情况。通常是将其蛋白产物的部分氨基酸进行测序。利用遗传密码(表 2-2),推断符合编码该蛋白的序列,可以化学合成一个短的单链 DNA 分子[14~30 个核苷酸长,称为**寡聚核苷酸**(oligonucleotide)]。由于遗传密码具有兼并性,通常不可能推断准确的序列,因此常合成包含所有可能性的寡聚核苷酸混合物。图 5-10 表示这种策略的一个例子。如果没有特异的 DNA 或蛋白质序列用以制备能靠杂交检测所需的 DNA 探针,必须用替代的方法。在过去的几年中人们开发了许多用于克隆 cDNA 的巧妙方法,大多数方法都纳入图 5-9 中所综合类别之一。首先如图 5-7 所示,用从所需类型细胞中制备的 mRNA 建立含有大量 cDNA 的集合体,并

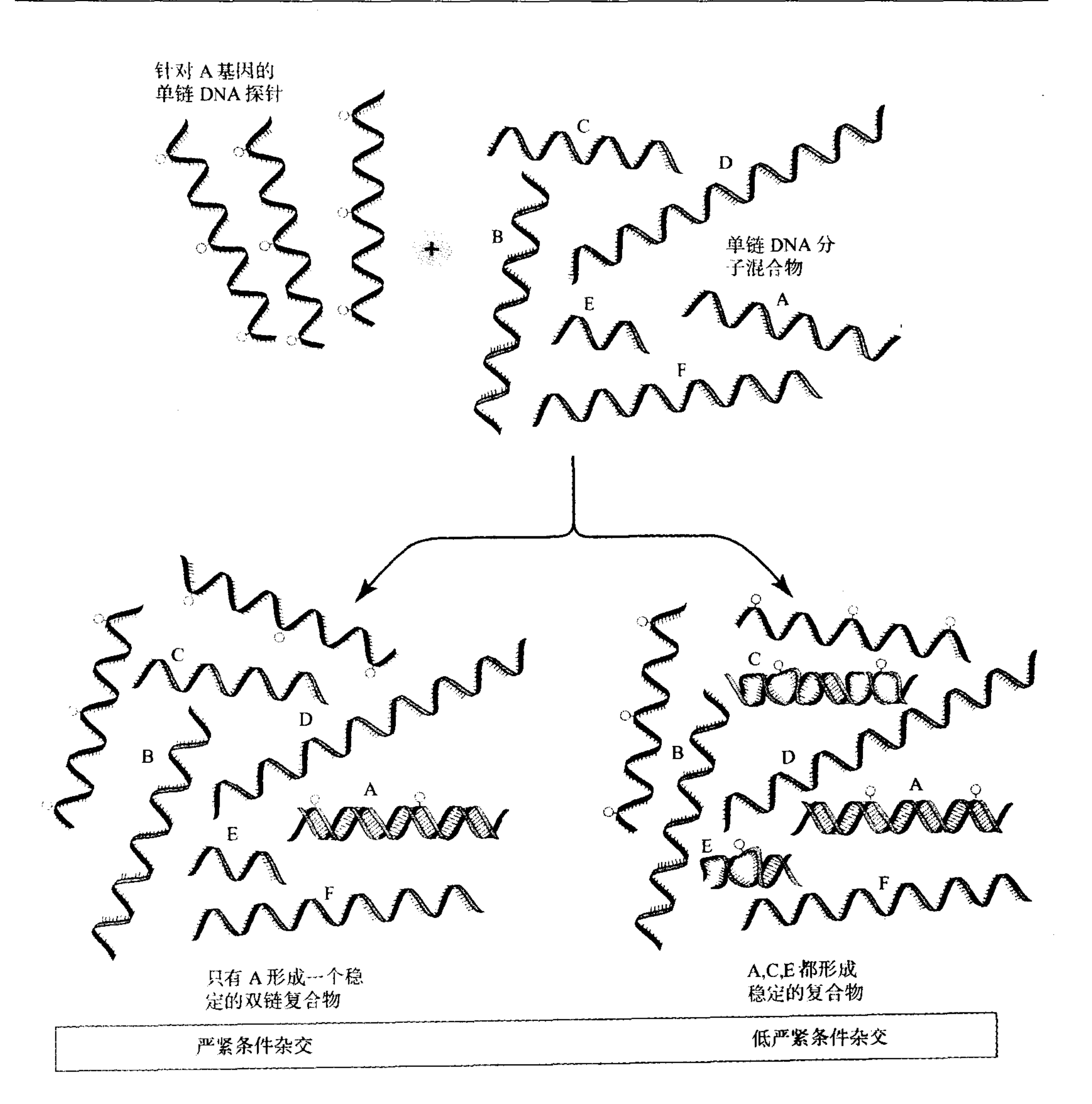


图5-8 核酸杂交。两个独立单链之间的配对过程以核酸固有的结构为基础,它有助于在互补DNA或RNA链之间的精确碱基配对。这种性质是在第二章中综述的DNA复制和转录基本过程的基础。在本例子中,标记A基因的单链DNA探针,以便检测其互补的序列。当与目标DNA分子的复杂混合物混合在一起后,该探针将只与和它精确互补的单链或密切相关的链形成稳定的双链。可调节杂交条件分离部分互补的双链,并只使完全配对的序列保持稳定的双链形式。这称为严紧条件杂交。在比较宽松的条件下,亦即低严紧条件下,密切相关但并不完全互补的单链仍然结合在一起。降低杂交条件的严紧度对于鉴定种属之间的同源基因非常有用。对于鉴定近亲基因家族成员也非常有用,因为这些成员的序列不相同但非常相似(引自 Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. Molecular Biology of the Cell. New York: Garland Publishing, 1994:306.)。

将它们连接或克隆到特殊的载体中,这个cDNA克隆的集合体就称为**cDNA文库**(cDNA library)。另一类统称为"表达克隆"的方法,是将cDNA连接到RNA转录和翻译所必须的序列上,使载体表达cDNA所编码的实际蛋白产物。通过用抗体检测cDNA编码的蛋白

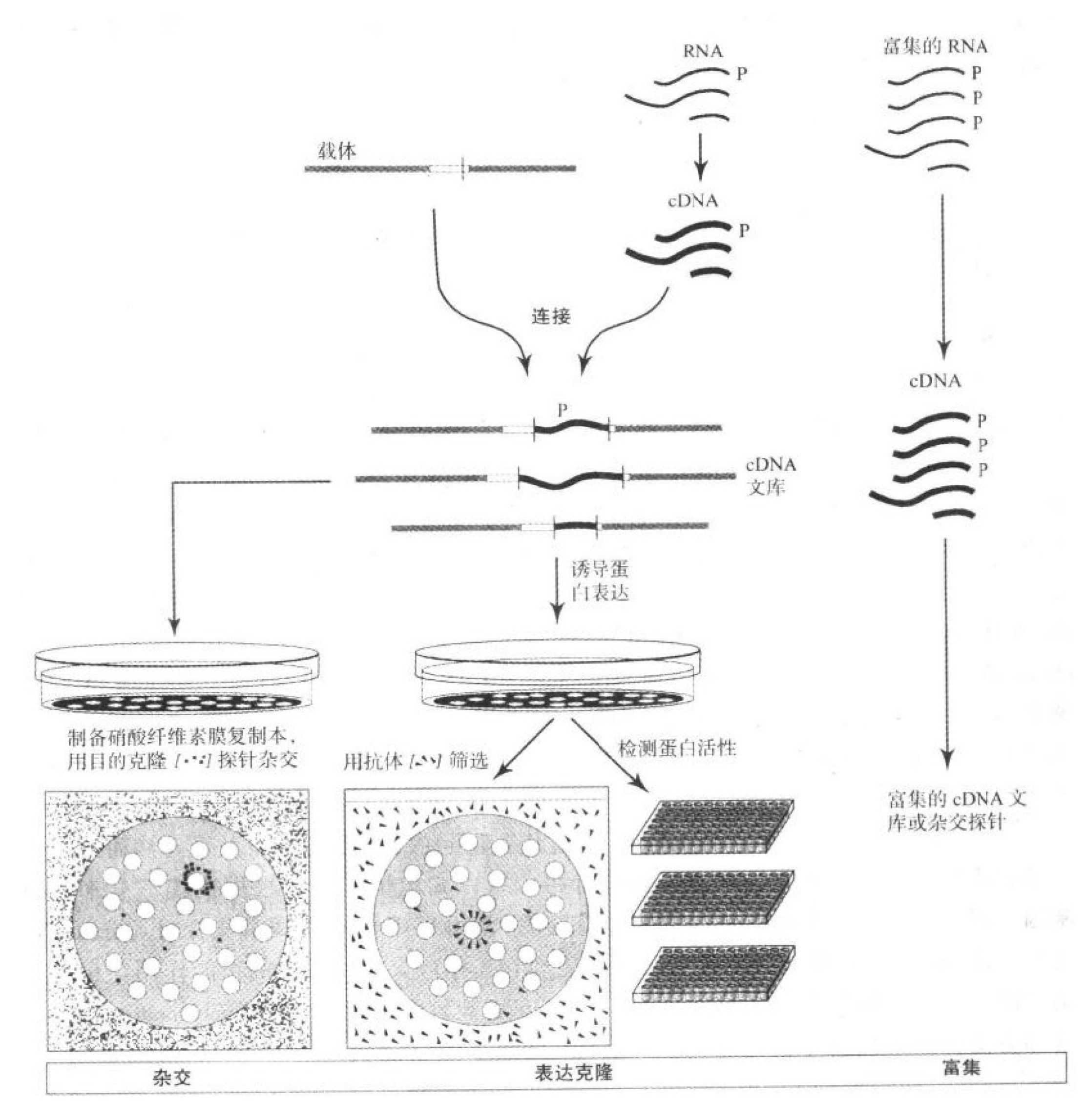


图 5-9 克隆一特异基因 cDNA 的方法。首先从表达(P)基因的组织中制备 cDNA,并将其插入克隆载体中(灰影线),然后将其转入细菌中(5-6 一样)。在产生的细菌群体中每一个细菌都带有一个含有不同插入 DNA 的载体。这个细菌群叫做一个"cDNA 文库"。如图左侧所示 cDNA 文库能被转到硝酸纤维素膜上,利用一个寡核苷酸或其他 DNA 探针来杂交筛选。另一种方法是利用"表达克隆"载体。该载体带有指导该克隆 cDNA 表达的序列(白框,图中部显示)。蛋白质产物有时与宿主细菌蛋白或其他蛋白融合在一起,可以用一个特异抗体(如果有的话),或用检测目标蛋白已知功能的方法来检测。利用硝酸纤维素膜或液体培养基中培养的单个菌落或菌落群可以筛选文库中大量的 cDNA 克隆。为简单计,图中只显示了若干克隆。实际上,经常筛选数十万到一百万个克隆。许多其他克隆技术依赖为所需特异序列富集 mRNA 或 cDNA 的方法(图中右侧显示)。富集方法有许多种,包括杂交筛选、按片段大小分离,和抗体筛选 mRNA - 核糖体复合物。消减杂交是一种特殊的富集方法。它是用相互杂交来除去两种细胞共有的序列,留下某一细胞特有的序列。富集的 mRNA 或 cDNA 能用来构建一个特殊的 cDNA 文库或制备杂交探针。

或利用该蛋白的功能就能检测目的 cDNA。还有一些进展就是用一些手段来富集杂交探针或富集 cDNA 文库本身得到相应于人们感兴趣克隆的序列。

氨基酸序列： —Phe—Trp—Met—Asp—Cys—Ala—

寡核苷酸序列： —TT$^{T}_{C}$—TGG—ATG—GA$^{T}_{A}$—TG$^{T}_{C}$—GC—

图 5-10 利用氨基酸序列信息构建 DNA 探针。如果一个蛋白的部分氨基酸序列是已知的，就可以合成一个寡核苷酸探针来与编码该蛋白的基因杂交，达到克隆该基因的目的。利用遗传密码表(见图 2-2)可以推测每个氨基酸的密码子；由于存在密码子的兼并性，许多氨基酸都有多于一个密码子，故实际上人们合成含有各种可能寡核苷酸序列的混合物。在上举例中，8 个 17-mer(2^3)的混合物可能包括正确的序列，并能克隆该基因。

用图 5-9 所列的方法之一一旦成功地克隆一个特定的 cDNA，就可以用许多方法来分析该 cDNA 本身。如测定其全长 DNA 序列、用遗传密码知识来推测其相应氨基酸序列。该 cDNA 也可用作杂交探针来鉴定相关的 DNA 或 RNA 序列，如 Southern 和 Northern 印迹。此外，cDNA 还可用作探针来克隆全长的结构基因，包括所有的外显子、内含子以及相邻的启动子序列。为了后一目的，如图 5-6 简述所示，需要用细胞总 DNA 构建基因组文库。实际上，通常用特殊化载体来加快大片段 DNA 的分离和大量克隆的快速筛选。利用 cDNA 作为探针来杂交可以鉴定含有目标基因组序列的重组克隆。通过基因组 DNA 和 cDNA 的序列之间的比较可以确定基因中内含子的位置(比较图 5-1 中成熟 mRNA 与完整基因)。基因组的和 cDNA 克隆之间的差别可能混淆。要记住，基因组 DNA 包括基因的所有成分，包括内含子、外显子，而来源于 mRNA 的 cDNA 仅包括被转录和加工的外显子。

过去，根据对蛋白质结构或功能的了解，从克隆 cDNA 着手简易而可行。然而近年来分子遗传学的进展已经使我们能够在没有相应蛋白知识情况下仅根据遗传信息来克隆基因。利用这种方法鉴定了一个基因组克隆后，就可以用该克隆做杂交探针筛选 cDNA 文库。此 cDNA 可以用来研究蛋白质的功能和它在相应人类疾病中的可能作用。仅根据某基因在基因组中的位置来鉴定一个基因的方法被统称为**定位克隆**(positional cloning)。这种方法将在第九章详细讨论。

(六) DNA 印迹法

一旦克隆了一个 DNA 片段，就能利用该片段作为杂交探针来鉴定个体间的多样性。其中一种技术就是 DNA **印迹法**(Southern blot)。其名字来源于它的发明者 Edwin M. Southern，与地理概念无关。在图 5-11 所示的步骤中，来源于某个体的基因组 DNA 被限制性酶消化。将由此产生的无数片段在电泳凝胶上按大小分开，然后将 DNA 转移到硝酸纤维素滤膜上并与标记探针杂交。只有与探针序列互补的那些片段被检测出来。当与正常图谱比较时，能发现以下两种突变差异，(a)任何大于 50 ~ 100bp 的重排，如缺失或插入，都能由于片段大小改变而被检出(图 5-12)；(b) 导致消化 DNA 限制酶的酶切位点产生或破坏的单个碱基差异会导致谱带变化(图 5-13)。图 5-12 和 5-13 描述了利用 DNA 印迹法检测这些突变差异类型(以Ⅷ因子基因为例)。Ⅷ因子基因的结构见图 5-14。

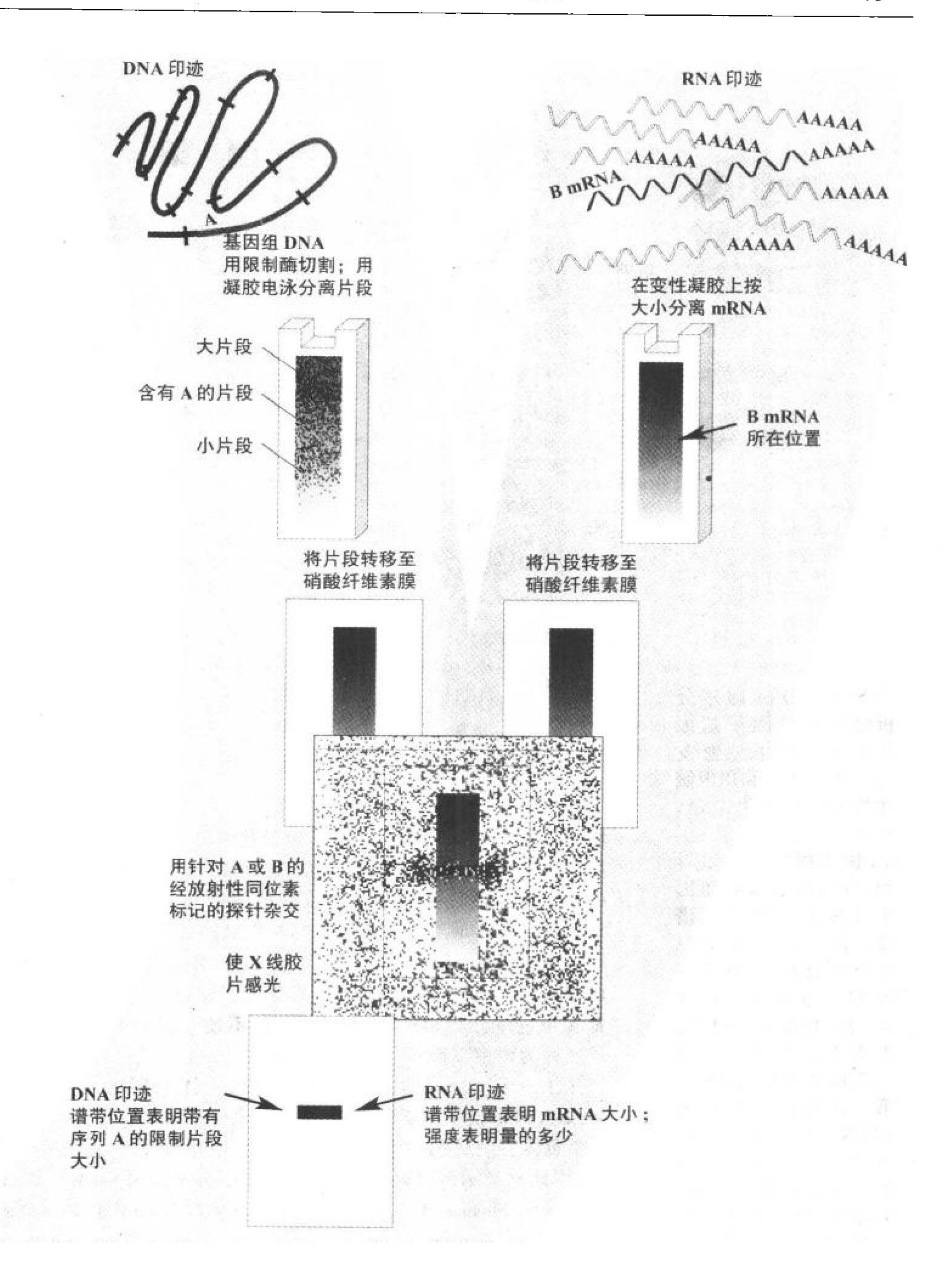


图 5-11 DNA 印迹和 RNA 印迹。这些技术能用来在复杂混合物中检测特异基因组 DNA 片段或特异 mRNA。DNA 印迹能在约一百万个基因组片段混合物中显示出一个特异的 DNA 片段(这里标为 *A* 基因)。从该目标片段在电泳凝胶中移动的距离能测出该片段精确的大小。RNA 印迹可以在源于一个组织样本的 10 000 或更多的 mRNA 复杂混合物中测出一个特异 mRNA 分子(这里标记为 B)。mRNA 的大小可以由它在电泳凝胶中移动的距离测出;其表达丰度可由电泳条带的密度测定。

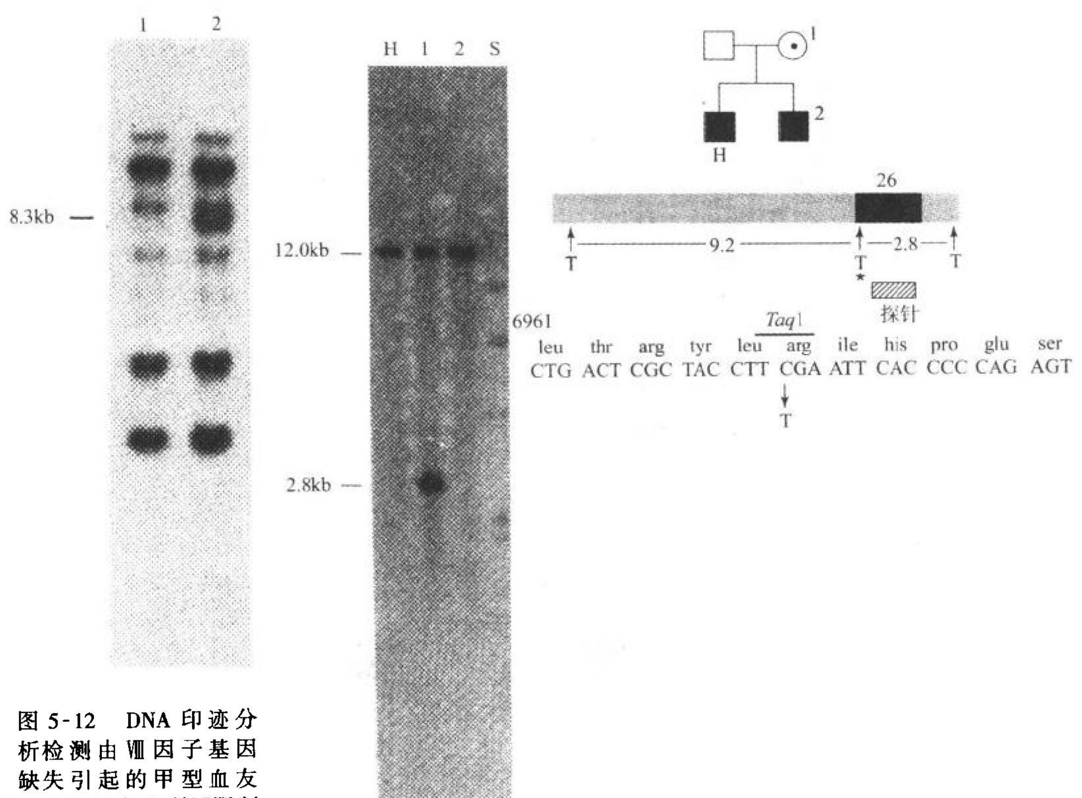


图 5-12　DNA 印迹分析检测由Ⅷ因子基因缺失引起的甲型血友病。DNA 印迹利用限制性酶 *Sst* Ⅰ 消化 DNA,并利用全长Ⅷ因子 cDNA 做为探针。可以看到 cDNA 测出来自Ⅷ因子许多不同区域的谱带。第一泳道的 DNA 由于在Ⅷ因子内没有第 23～25 外显子而导致 39kb 的缺失,使病人患甲型血友病。将第一泳道的 DNA 谱带与第二泳道内没有缺失的 DNA 谱带比较会发现缺少一条 8.3kb 的带。该谱带含有该 cDNA 探针测出的缺失外显子,这些外显子只在无缺失 DNA 样品中存在(引自 Gitschier J, Wood WE, Tuddenham EGD, Shuman MA, Goralka TM, Chen EY, Lawn RM. Detection and sequence of mutations in the factor Ⅷ gene of haemophiliacs. Nature 1985; 315:427～430.)。

图 5-13　用 DNA 印迹分析由Ⅷ因子基因单个碱基对突变引起的甲型血友病。图底列出突变区的 DNA 序列,显示单个碱基突变由 C→T。因为限制性酶 *Taq* Ⅰ 的识别序列为 TCGA(见图 5-5),由 TCGA到 TTGA 的变化导致该限制酶位点的丢失。在该基因图中特殊的 *Taq* Ⅰ 限制酶位点由一个 * 号标出,两侧相邻的 *Taq* Ⅰ 酶位点由 T 表示。DNA 印迹探针的位置也做了标记。这个探针能检出正常 X 染色体上 2.8kb 的 *Taq* Ⅰ 片段,不能检出正常的 9.2kb 的片段。因为该突变破坏了 * 号标记处的 *Taq* Ⅰ 位点,因此突变染色体产生了一个 12kb 的片段。该例中母亲(个体 1)同时带有一个拷贝导致 2.80kb 谱带的正常基因和一个拷贝导致 12kb 谱带的突变的 X 染色体。她的两个血友病儿子都从她遗传了该基因的异常拷贝,因此只显示出 12kb 的谱带(引自 Gitschier J, Wood WE, Tuddenham EGD, Shuman MA, Goralka TM, Chen EY, Lawn RM. Detection and sequence of mutations in the factor Ⅷ gene of haemophiliacs. Nature 1985; 315:427～430.)。

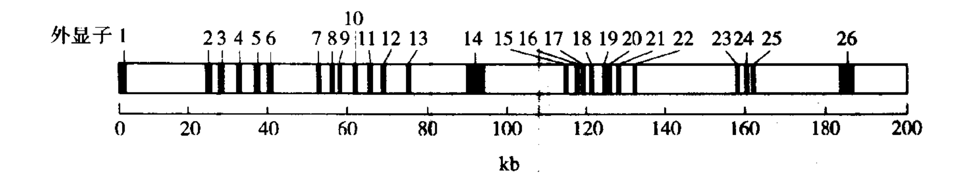


图 5-14 X 染色体上Ⅷ因子基因图。每个外显子都已标号。可以看出该基因非常大。Ⅷ因子基因缺陷产生与血液凝固有关的蛋白,导致血友病(经印到 Lawn RM, Vehar GA 的许可引自 The molecular basis of hemophilia, Sci Am 1986;254:48~54.)。

(七) RNA 印迹法

一个类似的杂交技术可以用来鉴定与标记探针互补的特定 mRNA。在这个被称为 RNA **印迹法**(Northern blot)的技术中,从特定的组织中制备 mRNA、依据大小在凝胶上分离、将其转移到滤膜上,再用 DNA 探针杂交(图 5-11)。实际上,DNA-RNA 杂交体比 DNA-DNA 杂交体更稳定。用该探针鉴定的带型表明该 mRNA 的大小,并证明相应的基因在该 mRNA 被提取的组织中转录("表达")了。杂交带的强度也大致代表了该基因表达的量。如果利用它来筛查突变,就能发现异常大小的 mRNA 以及表达水平的变化(图 5-15)。

(八) 聚合酶链式反应(PCR)

DNA 技术的一个对基础研究和遗传学诊断有深远意义的进展就是人们能够从复杂混合物中扩增出一个短的 DNA 片段。它使我们能在没有克隆出一个特定 DNA 片段之前就对它进行研究,从而大大地增加了分析速度。图 5-16 图解了该方法的原理。我们必需在已知该区域序列的基础上用化学合成的方法合成 2 个短核苷酸引物(通常 15~30 个核苷酸长,图 5-16 中 A 和 B)。引物加入到双链 DNA 中,随后加热使之变性("熔解")为单链 DNA。因为引物过量存在,每个引物都将在总 DNA 混合物中"找到"它的互补序列,并与之杂交。把一种能耐受 95℃高温的酶,即 *Taq* 聚合酶,加到该混合物中,该酶利用原始 DNA 为模板从复性的模板末端合成一条 DNA 链。该混合物又一次变性,剩余的引物变性,聚合酶反应又一次进行。整个过程很容易循环进行,每一循环仅需几分钟,完成 30~40 个循环不需要添加新酶或引物。如图所示,结果在两引物之间的 DNA 量呈指数增加。每次循环产物增加一倍,因此,30 循环后,原来的靶序列被扩增 2^{30} 倍,或大约 10 亿倍。这种变性、复性和聚合酶拷贝反应的反复循环,是在一热循环仪中按特需的温度和特定时间编程自动完成的。循环的实际结果是:我们可从纳克量的 DNA 开始进行 PCR 后,样品在凝胶上泳动,最后能观察到恰好相应于 A 和 B 引物之间长度(包括引物本身长度)的特异的 DNA 带。

PCR 也能扩增特定的 RNA 序列。首先,相同于图 5-7 所示的双链 cDNA 合成的第一步,首先 RNA 必需被反转录酶拷贝成 cDNA。在 PCR 的第一个循环中,单链 cDNA 被其中的一个引物拷贝成双链 cDNA,然后,再由 PCR 扩增,其方式如图 5-16 所示的双链 cDNA。

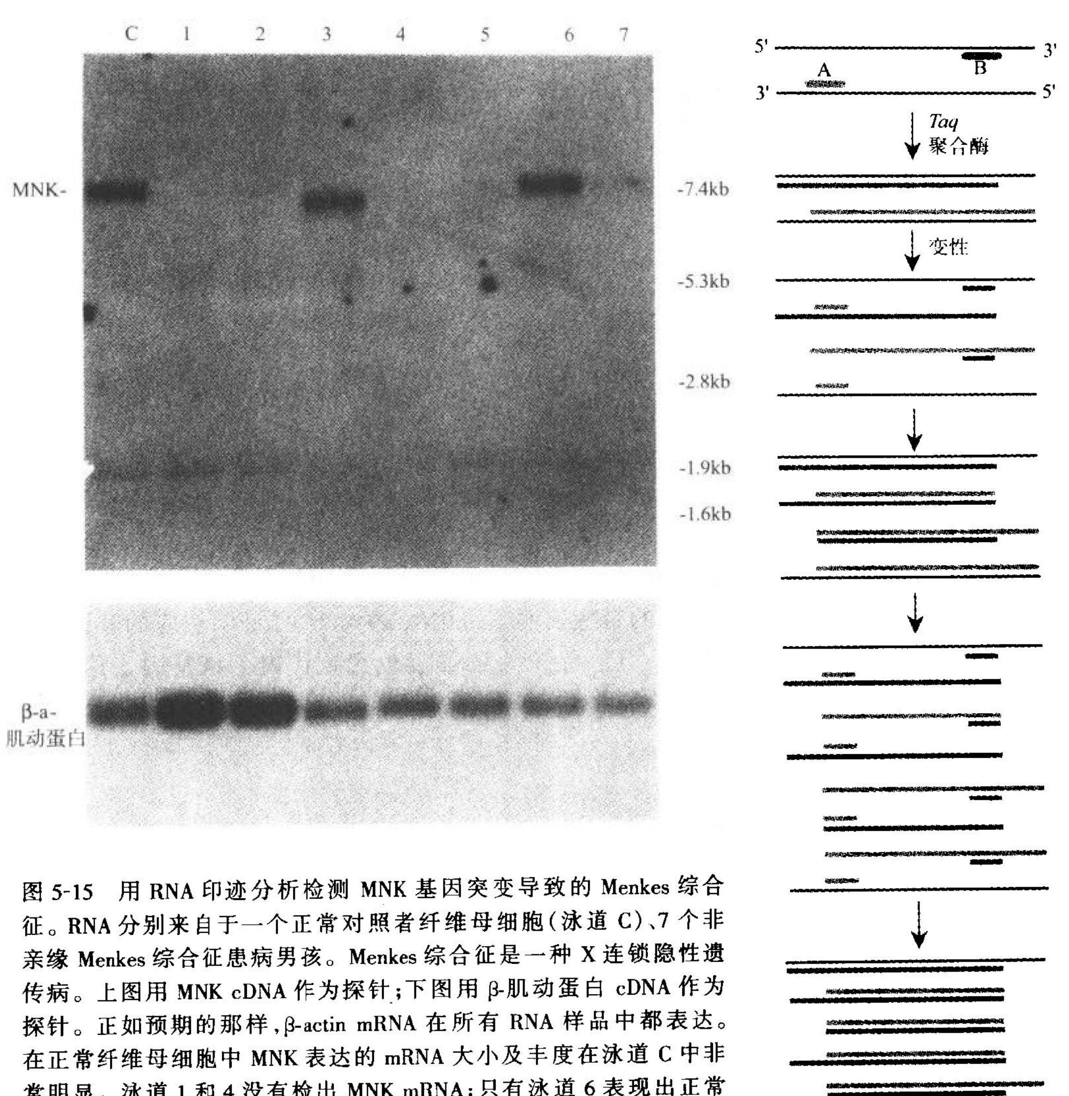


图 5-15　用 RNA 印迹分析检测 MNK 基因突变导致的 Menkes 综合征。RNA 分别来自于一个正常对照者纤维母细胞(泳道 C)、7 个非亲缘 Menkes 综合征患病男孩。Menkes 综合征是一种 X 连锁隐性遗传病。上图用 MNK cDNA 作为探针；下图用 β-肌动蛋白 cDNA 作为探针。正如预期的那样，β-actin mRNA 在所有 RNA 样品中都表达。在正常纤维母细胞中 MNK 表达的 mRNA 大小及丰度在泳道 C 中非常明显。泳道 1 和 4 没有检出 MNK mRNA；只有泳道 6 表现出正常大小及丰度的 MNK mRNA。虽然泳道 3 的 MNK mRNA 丰度看起来正常，但却比预期的小；泳道 7 的 mRNA 大小正常，但其量显著减少。泳道 2 和 5 的样品分别含一个正常大小的 mRNA 和一个大的 mRNA，但二者的量都显著减少(引自 Vulpe CB, Levinson S, Whitney S, Packamn S, Gitschier. Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. Nat Genet 1993; 3:7-13.)。

图 5-16　聚合酶链式反应示意图。利用 A 和 B 引物重复的 DNA 合成循环反应可以使位于二引物之间的 DNA 数量呈指数增长。分析大量循环(没显示)表明经过 20～30 次循环后主要产物是 DNA，它恰好始于 A 引物 5′端首，止于 B 引物 3′端末，见该图最后一组第三和第六个双链。

从基因组 DNA、mRNA 或任何其他的 DNA 扩增的 PCR 产物可用各种方法进行分析。扩增的 DNA 可作为特异探针用于上述的 RNA 印迹和 DNA 印迹分析;或用下述各种方法分析其特异 DNA 序列内容。PCR 是重组 DNA 技术的一次革命。鉴定人类遗传疾病缺陷的过程从过去的数周到现在大约几分钟到几小时就可完成。

(九) DNA 测序

现能容易地测定 PCR 产物或 DNA 克隆片段的精确核苷酸序列。最常用的方法之一是 Sanger 方法,它是一种酶促反应技术(图 5-17)。测定的结果是梯状带型,从带谱中就能

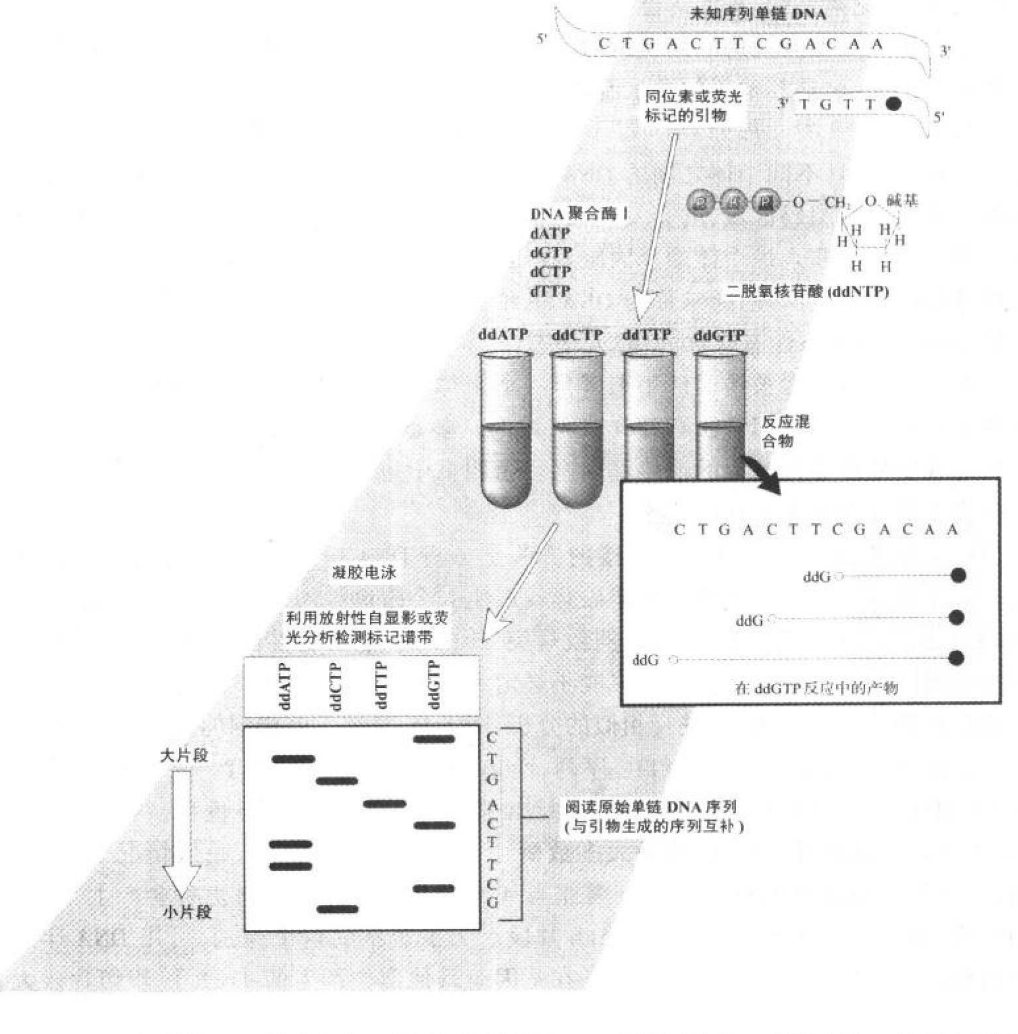


图 5-17 Sanger 发明的 DNA 测序方法。利用与待测单链 DNA 同源互补的一个引物和 DNA 聚合酶来进行互补链的合成反应。在四个反应体系中分别加入少量四种双脱氧核苷酸(ddATP, ddCTP, ddTTP, ddGTP)的一种,当一种双脱氧核苷酸被掺入不断延长的链时,因为没有羟基而导致合成反应终止。利用聚丙烯酰胺凝胶电泳分别分离上述四个反应体系中标记的合成片段,再通过 X 线放射自显影或荧光分析,就可以直接读出序列。

直接读出 DNA 序列。一次测序反应通常能测定大约 500 bp 长的片段。现在自动 DNA 测序仪利用上述方法能快速大量地测定样品。将含有 DNA 模板的标准酶反应体系放入测序仪中,仪器能自动分离和分析 DNA 片段,并将分析结果存入一个计算机磁盘文件中。该技术的不断进展已经使得测定人类全部基因组序列成为一个实际目标,并导致了人类基因组计划的建立(见第十章)。

(十) 多态性

分子遗传学技术的另一个领域是分析共同的 DNA 序列变化或多态性,它对人类疾病基因的鉴定(见第九章)有深远意义。开始人们对多态性现象只是感到有些好奇,但是当我们看到它在人类疾病的应用实例后,就会明白这些工具为我们提供了机会去跟踪人类家族疾病基因,这是其他任何方法都做不到的。

如前所述,人类大部分 DNA 都不是编码序列,因此就不会受到严格的选择。可能就是由于这个原因,不同个体之间的 DNA 差异并不少见。具体地说,如果我们察看十个个体的常染色体非编码区域,就会发现大约每 200~400 个核苷酸就有一个核苷酸变异。通过比较任意两个非亲缘个体的基因组,我们会发现无数的 DNA 序列差异。这些变异经常以简单的碱基变化出现,通常称为 DNA 序列多态性。独有的或局限于个体的小人群或家族的 DNA 序列多态性有时称为"私人多态性"。

多态性这一词常被错误地用来区分一个个体或家族中没有引起疾病的 DNA 序列变化和致病突变。在第四章已有较详细的讨论。多态性一词应该专用于至少频率为 0.01 的稀有或变异的等位基因的遗传基因座。第四章中也讲到,一些疾病突变如镰状红细胞疾病突变能够达到多态性的频率。

虽然检测 DNA 序列变化最直接的方法是分析 DNA 序列,但此法慢、昂贵,通常不适用于大规模筛查。有些序列变化能破坏或产生一个限制性酶切位点,因此易被 DNA 印迹或 PCR 方法测出。图 5-13 表示一例致病突变导致一个限制性酶切位点丢失,并用 DNA 印迹检测出来。一个与人类疾病突变不必定相关的共同 DNA 序列变化导致一个限制性酶切位点的产生或丢失,可进行相似的分析,图 5-18 列举了更详细的例证。在此例子中,A 染色体在 * 号部位含有 GAATTC 序列,而 B 染色体含有 GAATTT。限制酶 *Eco*R Ⅰ 能识别 GAATTC 序列,因此将在此部位切断 A 染色体而不是 B 染色体。该图也显示了 *Eco*R Ⅰ最邻近的酶切位点;因为大多数酶切位点不具有多态性,一般情况下,侧面位点固定,且各染色体之间没有变化。因此当 A 染色体被 *Eco*R Ⅰ 酶切后会产生 2kb 和 3kb 的片段,而 B 染色体将会产生单个 5kb 片段。为了观察到这个差别,可用 DNA 印迹来检测这些特异基因组片段。这些片段是在基因组其他部位产生的 *Eco*R Ⅰ 酶切片段大背景下来观察的。图 5-18 显示利用不同探针如探针 1、2 和 3 所产生的不同带谱。注意不同的带谱取决于哪些片段与标记探针重叠。(即使一个基因组片段的部分 DNA 序列与探针重叠,它也能检测出来,因为它们仍然能杂交;整个探针序列与基因组片段重叠是没有必要的)。该图显示了一个人有两个 A 染色体,一个 A 和一个 B 染色体,或两个 B 染色体可能获得的几种分析结果。这种导致特殊限制酶切片段长度变化的 DNA 序列多态性通常

称为**限制性片段长度多态性**(restriction fragment length polymorphisms, RFLPs)。如果已知多态性位点周围的 DNA 序列,如图 5-18 所示,通常用 PCR 分析 RFLP,较易而快。在此分析中用多态位点侧面的引物进行 PCR 扩增,产物用相应的限制性酶消化,就可以用凝胶电泳分析酶切片段的大小。

在解释 RFLP 分析时应记住三项关键事件:(a)通常被限制性酶检出从而导致了 RFLP 的碱基改变本身并不是致病原因,而是无功能后果的一种中性改变。然而它能用来标记特定的染色体从而跟踪它的遗传。除非有目的特殊选择(如图 5-13 举例),否则限制性酶消化几乎不能检出致病突变。(b)RFLP 等位基因的遗传严格遵守孟德尔式规律。因此,AA 和 BB 个体的所有后代可能都是 AB。(c)在利用 RFLP 跟踪一个染色体在一个家族中的遗传时,最有用的将是在大多数个体中均为杂合性的那些 RFLP(也称为标记)。其中原因在我们看到某些例子时才能完全明了,有一点也许是直观明显的,人们不能从纯合性 AA 中跟踪这两个特异染色体的遗传。

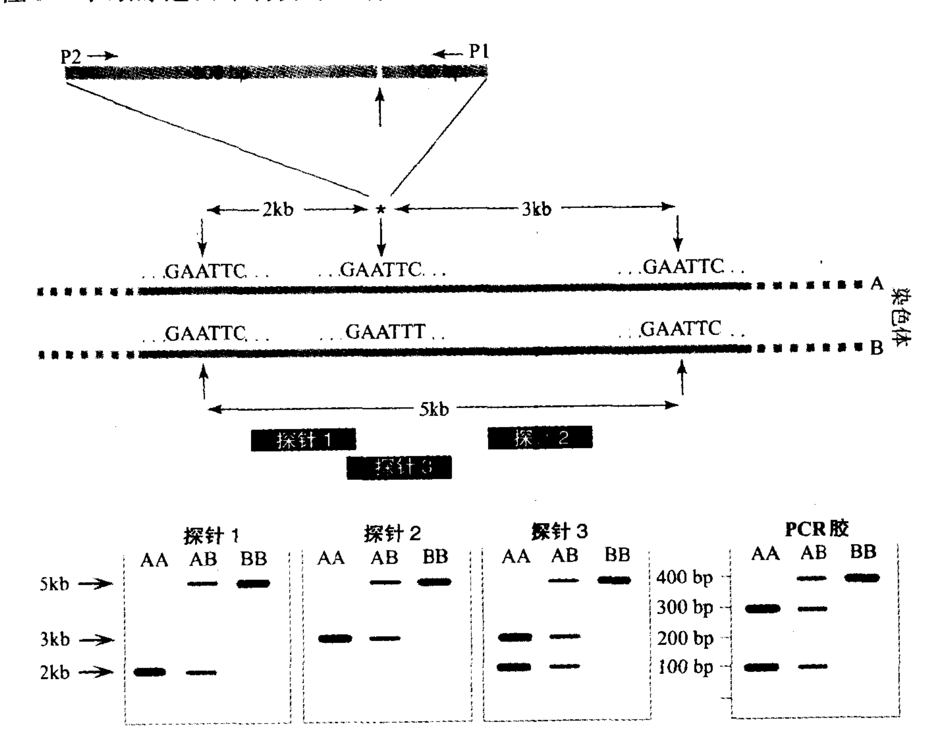


图 5-18 用 DNA 印迹和 PCR 分析限制性片段长度多态性(RFLP)。染色体 A 和 B 在标有 * 号处的序列有所区别,该区别导致染色体 A 存在 *Eco*R Ⅰ 位点,而染色体 B 则没有。*Eco*R Ⅰ 识别 GAATTC 序列。PCR 引物 P1 和 P2 是专门设计来扩增该 *Eco*R Ⅰ 多态位点周围基因组 DNA 的。用限制酶 *Eco*R Ⅰ 处理该基因组 DNA。请注意在图底部的 DNA 印迹图根据所用探针的不同而有所变化,因为只有当一个片段与所用探针至少部分重叠时该片段才能被检测出来。最右侧的图显示用限制酶 *Eco*R Ⅰ 切割引物 1 和 2 扩增的 PCR 产物后的分析结果。AA 是染色体 A 的纯合子;AB 是杂合子;BB 是染色体 B 的纯合子。

到目前为止我们只讨论了产生或破坏一个限制性位点的一个核苷酸变化的 RFLP。图 5-19 显示了一种比较少见但非常有用的多态类型。这种多态性称为长度多态性、小卫星或"可变数目串联重复序列"(VNTR)。它们是由一些头尾相连的短序列以串联方式重复多次形成的。这些序列如果有功能现在尚不清楚,但一个元件中的重复单位数量(即元件的总长度)在同源染色体中经常变化。在这种情况下,任何限制性酶切断这个串联重复序列的两侧将会产生一个片段,而其长度会反映出 VNTR 的大小。因为这个重复序列的总长度在不同染色体之间变化很大,可能存在很多数量的等位基因,这一特点使这些多态性非常有用。图 5-19 显示了 一例 VNTR 多态性。探针 1 对应于此特殊 VNTR 多态性独有的 DNA 片段,因此利用 DNA 印迹只能检测此单一位点。所有这三个个体在这个基因座都呈现杂合性。探针 2 与 VNTR 多态性本身的串联重复序列配对。因为同样的串联重复序列为许多散布于基因组中的不同 VNTR 多态性,探针 2 在 DNA 印迹中将从这些众多多态性中检测到许多带谱。利用探针 2 这样的重变性探针检测到的复杂谱带同时显示出许多多态基因座,因此在两个个体很少一样。这种带谱图经常被称为"DNA 指纹",如图 5-20 所示的一例。这种 DNA 同一性分析在亲子鉴定和法医学中已经广泛应用。然而实际上现在这类分析经常用类似图 5-19 中探针 1 的独特探针,每个探针只检测单个高度多态的 VNTR。因为这类 VNTR 有大量的等位基因,故两个非亲缘个体携带相同的两个等位基因的机率非常低,经常小于 1%。通过检测 4 个或更多这种独立的多态性,任何两个非亲缘个体在所有这几个基因座都共有恰好相同等位基因对的机率极小,除非他们是单卵双生子。图 5-21 是一个法医 DNA 分析的例子。

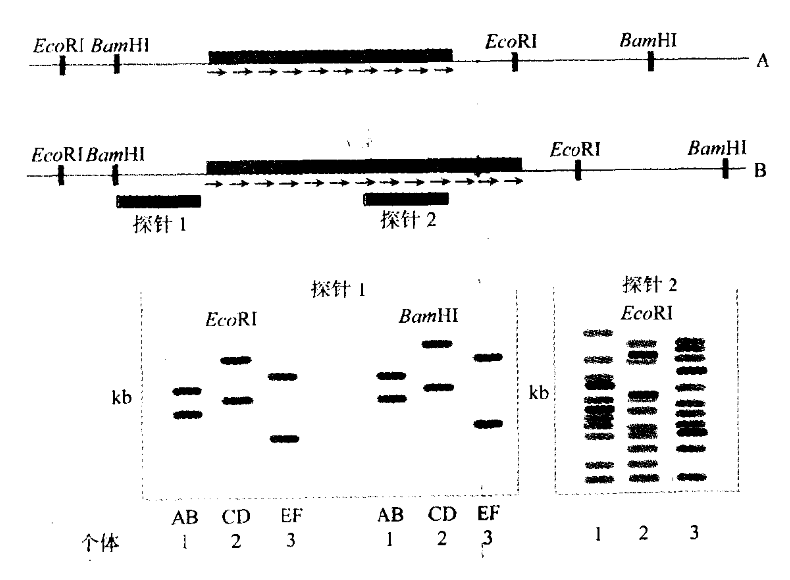


图 5-19　可变数目串联重复(VNTR)多态性。画框的区域是一长串短串联重复序列。不同染色体中重复元件的数量不同,可以遗传。利用 DNA 印迹可以观察出不同长度,而与用什么限制性酶几乎无关。在图中利用 *Eco*RⅠ和 *Bam*HⅠ两个限制性酶和探针 1 证明所有三个个体都是杂合子。这也是该类多态性的优点。除探针 1 检出独有的 VNTR 外;探针 2 对应于串联重复区域,检出许多相关的 VNTR,因此产生更为复杂和高度多态的图谱。

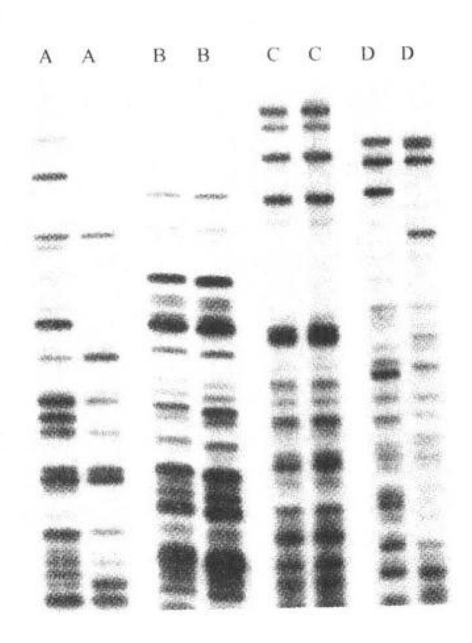
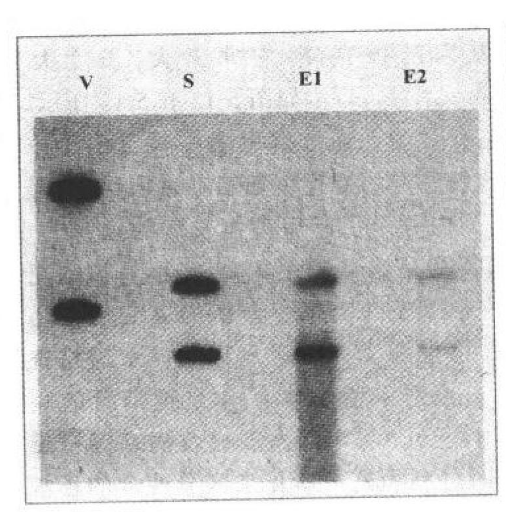


图 5-20 利用单个探针检测大量分散于基因组内 VNTR 多态性的 DNA 指纹分析。每对样品代表孪生儿的 DNA。B 和 C 是完全配对组,表明它们都是单卵双生子;A 和 D 是异卵双生子。注意异卵双生子之间有大约一半可变谱带相同,与一级亲属一样。



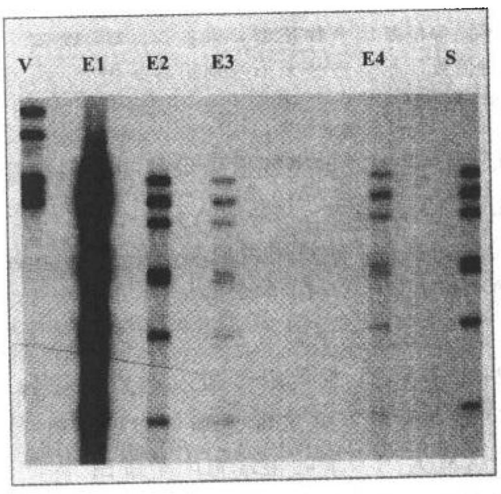


图 5-21 VNTR 多态性和法医学 DNA 分析。图左侧 DNA 印迹分析显示独特 DNA 探针检测单个 VNTR 位点,与图 5-19 中探针 1 相似。被害者(V)与嫌疑犯(S)的 DNA 显示了不同的谱带。从犯罪现场采集的物证制备的 DNA 与嫌疑犯的 DNA 谱带吻合。从各种 VNTR 等位基因在总人群中的频率就可估计出吻合嫌疑人的概率。通过对许多不同的 VNTR 多态性进行相似的分析,这种吻合的概率就会非常低。图右侧的 DNA 印迹分析是同时用了 4 个不同的单个位点特异 VNTR 探针。提取的物证与嫌疑犯之间的结果完全吻合。目前的法医学分析通常采用至少四种独立的、高度多态的 VNTR 标记(经 ICI Americas Inc. 的许可引自 DNA Fingerprinting SM and DNA Profiling. Germantown. MD: Cellmark Diagnostics. 1989. .)。

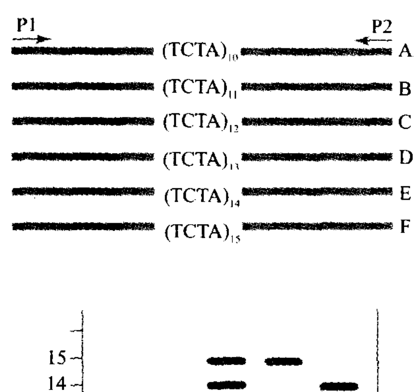
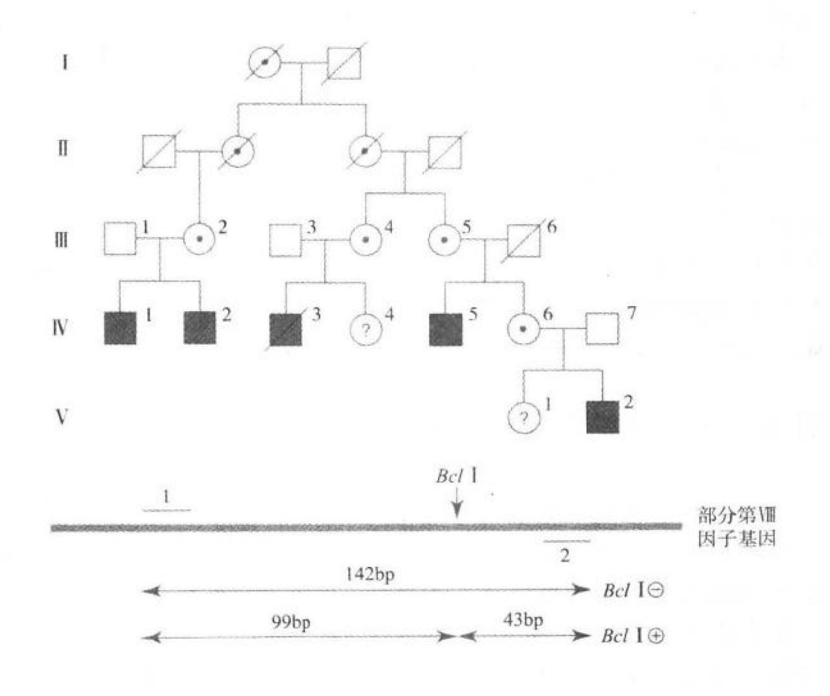


图 5-22 一个简单重复序列(SSR)多态性。该多态性的每一个等位基因都带有简单重复序列 TCTA 不同数量的拷贝。当用引物 1 和引物 2 来扩增该位点时,各种不同等位基因产物长度差别按该 4 碱基对倍数计算。这种大小差别很容易用凝胶分析解决。

因为大多数 VNTR 的重复区可长达数千 bp,标准 PCR 很难扩增到如此长度。因此 VNTR 通常必需用 DNA 印迹分析。然而最近开发出另一类可用 PCR 分析的高度多态性标记,因此更适合于大规模 DNA 分型。虽然这类多态性也含有串联重复序列,但基本的重复单位非常短,通常只有 2、3 或 4 个碱基对(分别为 2、3 和 4 核苷酸重复)。这类多态性被经常称为"微卫星"或简单序列重复(SSRs)(图 5-22)。正如第九章将要讨论的,在过去几年中鉴定的 10 000 个以上这类标记,显著地加速了人类及其他种类生物的基因制图和连锁研究,极大地促进了纯粹基于基因位置而进行的疾病基因鉴定(定位克隆)。最近,人们鉴定出位于一些基因内特异的三联体重复的长度不稳定变异,认为它是人类疾病突变的一个重要的新机制(见第三和第八章)。

DNA 序列多态性的有用价值在我们研究一个例子后会更为明了。正如第七章将要论述,甲型血友病是一个 X 连锁隐性疾病,由Ⅷ因子基因突变引起的。Ⅷ因子缺陷导致不能形成适当的血液凝块。Ⅷ因子基因已被克隆,它非常大,长度大约为 200 000bp(见图 5-14)。如图 5-12 和 5-13 所示,利用 DNA 印迹有时能在血友病患者中检出由于该基因中大片段缺失或点突变而导致的一种限制性酶切位点的变化。如第七章所讨论,最近利用 DNA 印迹分析一大群病人检出了Ⅷ因子基因内一个引起甲型血友病的独特的重排。然而大约有 1/2 导致甲型血友病的Ⅷ因子基因突变是错义或无义突变,不易被常规 DNA 印迹方法检出。在研究实验室之外,筛查大的Ⅷ因子基因每一个新的突变现在还不可能。图 5-23 显示甲型血友病一个典型家族,在第Ⅳ和Ⅴ代有患病男性。系谱检查发现Ⅲ-2、Ⅲ-4 和Ⅲ-5 是甲型血友病肯定杂合子。根据系谱个体Ⅳ-4 是携带者的风险为 50%,她在建立家庭之前想确切知道这种危险性。通常,人们发现血友病基因携带者血液中含有大约半量的正常Ⅷ因子蛋白,因此有时可利用这一所见来区分纯合的正常妇女和携带者。但是在正常和携带者之间有很大的重叠,所以利用蛋白定量的方法有时不能明确区分。这样的家庭中,确定Ⅳ-4 是否是携带者更准确的方法就是利用 DNA 序列多态性作为携带突变等位基因染色体的标记。图 5-23 中显示Ⅷ因子基因内一个内含子区域含有限制性酶 *Bcl* Ⅰ 多态性的限制图谱。

用位于多态性限制位点两侧的 PCR 引物 1 和引物 2 扩增靶区域。用 *Bcl* Ⅰ 切割扩增后的 DNA,切割后的 DNA 片段直接进行凝胶电泳以分析该位点是否存在。如果存在 *Bcl* Ⅰ 位点,142bp 的产物将被该酶切割为 99bp 和 43bp 两个片段;如果不存在以上位点,将会看到 142bp 的全长片段。所有血友病男性患者都有 142bp 片段,说明带有血友病突



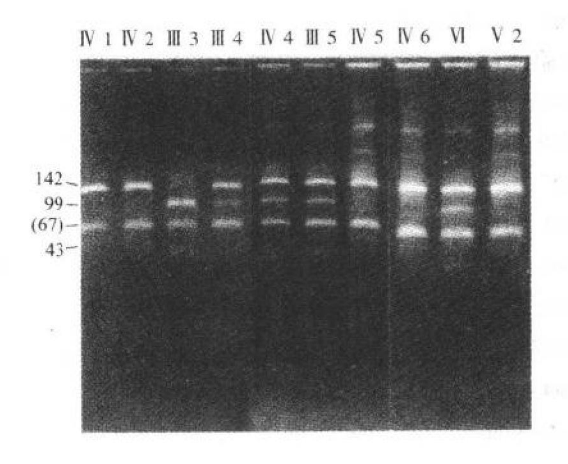


图 5-23 用 DNA 序列多态性分析甲型血友病家族。在系谱的下面是 PCR 扩增片段所在区域示意图,显示了引物以及 *Bcl* Ⅰ 多态位点所在位置。利用引物 1 和 2 进行 PCR 会扩增出 142bp 的产物,如果有 *Bcl* Ⅰ 位点存在,用 *Bcl* Ⅰ 酶时将分别产生一个 99bp 和一个 43bp 的片段。照片显示的是实际分析结果,每个泳道上面的数字与系谱中的个体相对应。在 67bp 处有一个固定的不明来源的条带,在本次分析中可以忽略。在所有的男性患者中,甲型血友病基因都与 142bp 条带(即缺少 *Bcl* Ⅰ 位点)的出现有关。仔细观察会发现个体Ⅳ-4 一定是甲型血友病携带者。个体Ⅴ-1 由于她母亲是该多态性的纯合子而不能断定(经 Kogan SC, Doherty M, Gitschier J. 的许可引自 An improved method for prenatal diagnosis of genetic disease by analysis of amplified DNA sequences. Application to hemophilia A. N Engl J Med 1987;317:985 ~ 990.)。

变基因的X染色体无*Bcl* Ⅰ位点。个体Ⅳ-4从她母亲Ⅲ-4遗传了*Bcl* Ⅰ等位基因,因此她就是一个血友病的携带者。这个例子说明了上述的一个重要论点:*Bcl* Ⅰ RFLP本身(在一个内含子内)并不是血友病突变,它只是作为一个染色体标记;其他带有同样等位基因的染色体并不带有血友病基因,这可以在个体Ⅳ-6带有正常等位基因的例子中得到证明。在这个例子中我们讨论了X连锁隐性遗传病,这类分析也同样可成功地用于常染色体显性和隐性遗传病。在今后的各章中我们将会遇到许多这样例子。此外,这种基于PCR技术的分析非常快速,几小时就能得到结果,而且只需要每个个体微量的基因组DNA。例如PCR能利用唾液样品(包含足够的上皮细胞为PCR提供足量的DNA)或外科病理实验室储存的久逝病人组织的石蜡包埋块来提取DNA进行分析。

(十一)检测DNA序列差异的其他方法

因为不是所有的DNA序列变化都产生限制酶识别位点便于检测其改变,因此人们开发了其他一些更通用的方法来筛查任何已知的序列改变。PCR有获取一个很纯的DNA片段样品的能力,因此可能进行精确的杂交实验检出单个核苷酸变化。图5-24描述利用这样方法检测导致镰状细胞贫血(见第六章)的突变。除了利用合成的寡核苷酸进行PCR以外,人们还制备对应于该片段内部跨越DNA序列差异部位的**等位基因特异寡核苷酸**(allele-specific oligonucleotides,ASO)。ASO通常长15~19bp,主要通过控制温度和盐浓度进行精确的杂交实验能够使"完全配对"的ASO有效地与靶片段杂交;而带有一个错配核苷酸的ASO将不会杂交。一般情况下,人们合成针对某一特定区域两种可能序列的ASO,标记这些探针后,在PCR扩增产物上分别对这两种探针进行实验,分析它们是否与PCR扩增的DNA杂交。

图5-25显示ASO用于图5-23同一血友病家族检测的过程。已合成一个与含有*Bcl* Ⅰ识别位点(TGATCA)染色体精确配对的ASO。另一个ASO与不含*Bcl* Ⅰ酶切位点、而含TGAACA序列的染色体精确配对。用PCR扩增取自此家族的基因组DNA,然后把少量扩增后的DNA加到椭圆槽内的硝酸纤维素膜中。上槽用标记的*Bcl* Ⅰ⊕的ASO杂交;下槽用*Bcl* Ⅰ⊖的ASO杂交。洗去多余的ASO探针后,用滤膜使X光片感光。结果与图5-17显示的凝胶分析结果完全一致。但是请注意,这种ASO方法不依赖染色体之间核苷酸差异部位是否存在限制性酶位点,它可用于检测任何部位差异。当然,我们必须知道关于一个特定区域的序列信息来进行分析,因为寡核苷酸PCR引物和ASO依赖这些信息。现已开发很多其他基于PCR的方法来快速检测已知的序列变化。其中一些方法特别适于自动化分析,很可能成为将来临床实验室中广泛应用的DNA检测方法。

对精确突变不明的病例,人们开发了一些扫描靶基因大片段的方法用以寻找可能的序列差异。即使是大如Ⅷ因子基因这样的基因(图5-14),现在也可用PCR扩增所有编码的外显子,并有效地筛查这些序列在特定病人DNA中独有的突变。图5-26描述了其中一些技术。**变性梯度凝胶电泳**(denaturing gradient gel electrophoresis,DGGE)利用带有正常和突变序列的DNA片段在一种特殊凝胶中移动速度不同的性质。**单链构象多态性**(single-strand conformational polymorphism,SSCP)分析是一个更简单的方法。它先使PCR扩

正常 DNA

PCR

5' GACTCCTG A GGAGAAGT 3'
3' CTGAGGAC T CCTCTTCA 5'

镰状细胞贫血病人 (SCA)DNA

PCR

5' GACTCCTG T GGAGAAGT 3'
3' CTGAGGAC A CCTCTTCA 5'

变性，加到双份硝酸纤维素膜“点杂交”

正常 DNA　SCA DNA

与标记的正常 ASO 杂交

5' GACTCCTG A GGAGAAGT 3'

与标记的 SCA ASO 杂交

5' GACTCCTG T GGAGAAGT 3'

正常 DNA　SCA DNA

正常 DNA　SCA DNA

图 5-24　利用等位基因特异寡核苷酸(ASO)检测 DNA 中单碱基差别。在本例中,ASO 用于检测正常 DNA 和含有镰状细胞贫血致病突变 DNA 之间的差别。此病是由 β 珠蛋白基因中 A 转变成 T 造成的(见第六章)。在严格的杂交条件下,标记的“完全配对”ASO 将与目标 DNA 结合,而含有一个错配碱基的 ASO 将不会结合;这就能明确区分开两个 DNA 样品。

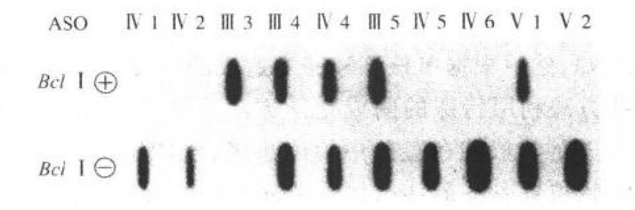


图 5-25　利用 ASO 检测 PCR 扩增的 DNA 分析图 5-23 的系谱。分析结果与图 5-23 凝胶分析结果相一致(经 Kogan SC, Doherty M, Gitschier J. 的许可而引自 An improved method for prenatal diagnosis of genetic disease by analysis of amplified DNA sequences. Application to hemophilia A. N Engl J Med 1987;317:985～990.)。

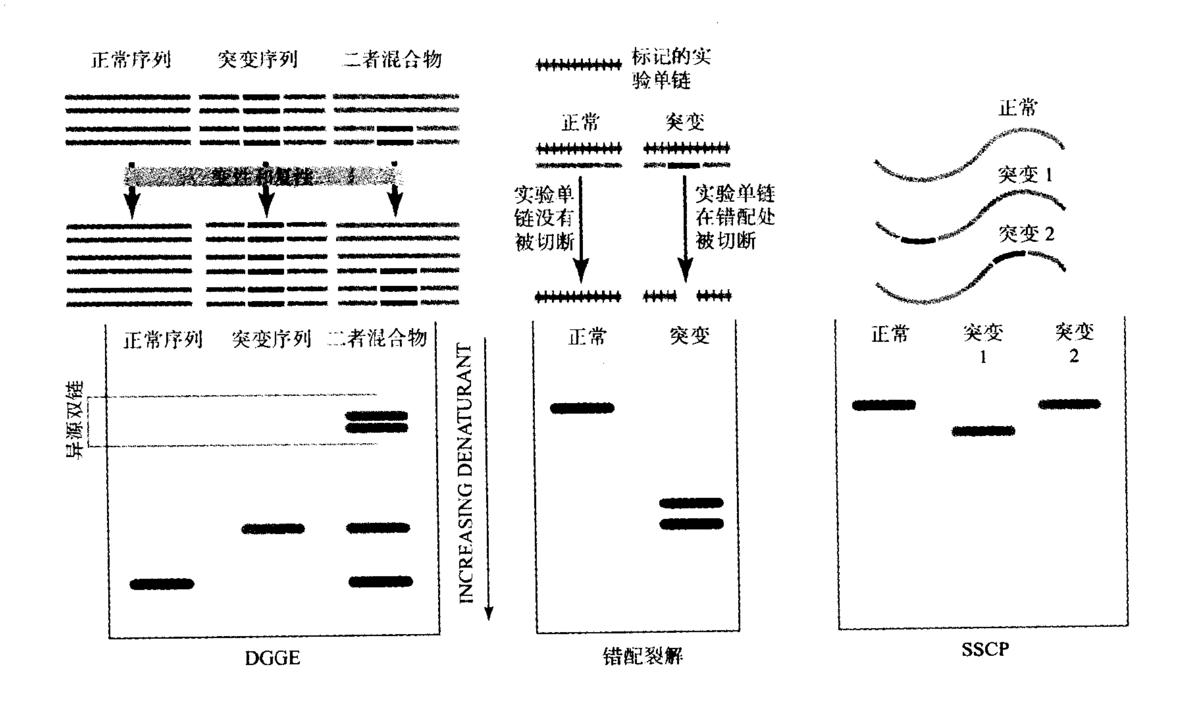


图 5-26 突变检测的筛查方法。用图解的方式描述了通过筛查一个 DNA 序列来鉴定新突变的几种方法。变性梯度凝胶电泳(DGGE)是利用 DNA 两个互补单链的精确序列在解链行为中出现的细小差异。电泳胶从上到下形成了一个变性条件由弱到强的梯度。当一个双链 DNA 片段在胶中行进时,行进速度在该双链开始解链的位置会突然放慢。在正常和突变的 DNA 序列之间一个碱基的差别将会导致解链行为轻微的差别,进而导致在胶中行进速度的差别。正常和突变单链的杂合双链会有一个错配的位点,这样会导致其在较低的变性条件下解链进而显著地改变其行进速度。当设定条件检测异源双链突变时,能达到 100%的检出率。错配裂解方法在标记待测 DNA 或 RNA 单链和一个正常对照序列之间形成异源双链。突变将导致在相应的位置出现错配。然后可用各种化学或酶方法在错配的位置切断异源双链。标记的被切断或未被切断的待测单链可以通过变性凝胶观察到。相加起来达到待测单链全长小片段的存在即表明存在突变。另外,谱带的大小提供相应突变所在位置的信息。已经开发的化学错配裂解方法能测出几乎 100%的突变。单链构象多态性(SSCP)分析从非变性胶目标区域分离出单链 DNA。每个单链的构象因其精确序列长度不同而有轻微变化,这种构象差别可由凝胶电泳速率反映出来。然而并不是所有的序列变化都能导致速率的显著差别,因此只有一部分突变能用这种方法检测出来。

增的产物形成单链 DNA,然后根据单核苷酸突变造成的在凝胶中移动的速度差异来区分它们。最后还有一些方法利用特殊的酶或化学方法在突变和正常序列之间的错配位点有效地切割 DNA。切开的片段能在凝胶中检测出来:既能鉴定出突变的存在,也能鉴定出它在该片段中的大致位置。在最佳条件下,SSCP 方法在一个特定的 DNA 片段中能检测出高达 70%~80%的序列差异,而 DGGE 和一些错配切断方法能检测出将近 100%的序列差异。这些突变筛查技术的改进终将导致更广泛地常规诊断人类遗传病的特异突变。

(十二)基因转移

基因序列特别有用,因为用它能精确地推测编码蛋白的氨基酸序列。但是不可能察看基因序列或它的侧翼区域就推测这个基因是如何被调节的或在什么组织中表达。若想获得这些信息,最好有一个供基因转移的系统。在该系统中,能将某基因测序、依照人们意愿的方式除去某些序列或改变其他序列,然后将它放回细胞内研究其表达为 mRNA 和蛋白质的情况。基因可以用各种方法送入培养瓶中培养的细胞内。这些方法包括 DNA 与磷酸钙简单共沉淀于细胞表面;或利用电休克(**电穿孔**,electroporation)刺激细胞摄取 DNA 或与脂质共孵育(**脂质体转染**,lipofection)等。用这些方法可以诱导许多不同类型的细胞摄入 DNA,而后 DNA 进入细胞核并被表达。在少部分这样转染的细胞中,外源 DNA 确实随机地共价整合到染色体中。一些重组病毒载体如反转录病毒和腺病毒都能用来有效地介导外源基因进入细胞。第十三章将更详细地讨论这些方法。将重组基因转入其他细胞的技术既能生产正常或改变了氨基酸序列的蛋白质,又能研究有调节基因表达作用的 DNA 序列。在人类遗传学中,这些技术用于生产大量的治疗遗传缺陷患者的重组蛋白质,如生产重组Ⅷ因子治疗甲型血友病。突变 DNA 序列功能的分析对区分真正的致病突变和中性突变以及了解疾病的分子发病机制至关重要。最后,将外源基因导入人类细胞治疗疾病是基因治疗新领域的目标(第十三章)。

转基因(transgenic)小鼠技术为在整个动物体内操作基因提供了一组有力工具。标准的方法是,用很纤细的移液管直视下将纯化的 DNA 注入小鼠刚受精后的卵细胞中。再将该卵细胞植入不同的小鼠中使其开始正常的胚胎发生过程。大约 10%~30%这样小鼠后代中会带有注入 DNA 的若干拷贝,并共价整合到生殖细胞 DNA 中。因此这些序列将以孟德尔方式遗传给后代。在很多情况下,注入的基因也被鼠转录。这给我们提供了一个极好的方法来研究指导某一特定基因正常发育和组织特异性表达必需的序列。图 5-27 描绘了这个过程。图 5-28 是这种试验所产生的小鼠照片。该试验注射的是按此方法设计的生长激素基因,使小鼠产生大量生长激素。该鼠与同窝正常小鼠相比体型明显增大。

近来一个令人振奋的进展是人们能对小鼠基因组进行更精确的定向操作。在这些方法中,DNA 被导入组织培养的胚胎干细胞(ES)中(图 5-27)。ES 细胞在组织培养中能被修饰,甚至经过几次传代后仍保持参与发育中的小鼠胚胎遗传组成的潜能。这样,当 ES 细胞用复杂的生物工程技术在体外成功地对某一个特定基因敲除或精细地改变部分序列后,由这些 ES 细胞产生的小鼠也会带有这一基因缺陷。将带有预定变化的 ES 细胞注射入囊胚中,并将这个早期小鼠胚胎重新植入受体鼠。由这样一个注射囊胚产生的小鼠将是由来源于原始囊胚细胞和 ES 细胞组成的混合体。一个由不同来源细胞构成的小鼠称为**嵌合体**(chimera)。在一个嵌合体小鼠中从 ES 细胞来源的生殖细胞(卵子或精子)将会稳定地携带最初导入其中的遗传改变,使之能构建稳定的小鼠品系。如果 ES 细胞中一个特定的基因被破坏,该嵌合体小鼠的后代之间的交配最终会产生该基因缺陷的纯合小鼠。利用这种方法产生的一个基因被破坏的小鼠经常被称为“**剔除**”(knockout)小鼠。

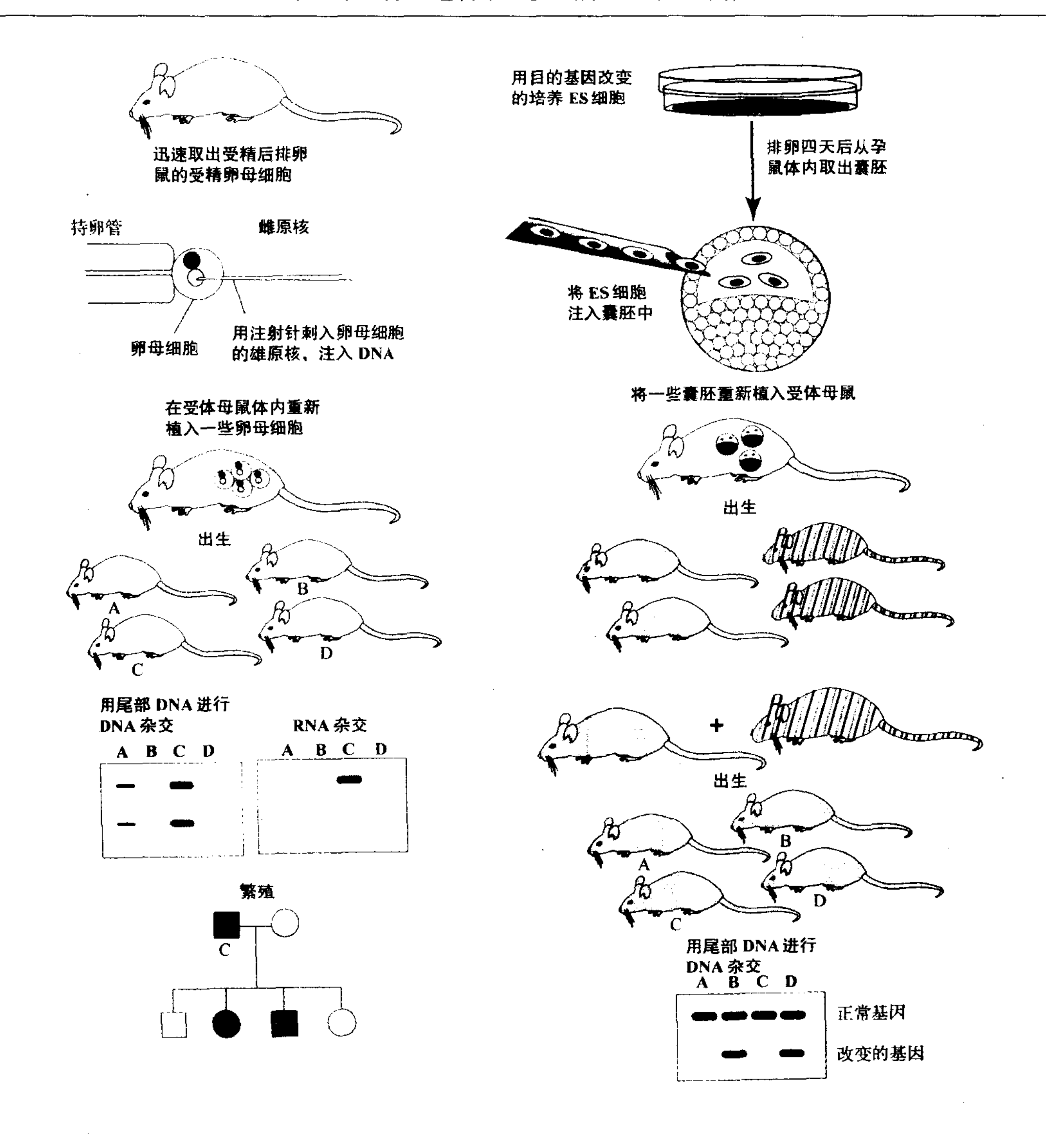


图 5-27 转基因小鼠的制备。可以用原核注射方法制备转基因小鼠。用手术方法在受精后和雄原核与雌原核融合之前取出卵子。再将待转移到小鼠基因组内的克隆 DNA 注射到雄原核内；雌雄原核融合后，注射的 DNA 在一部分卵子中就会随机地共价整合人小鼠基因组位点。用手术方法重新将这些卵子植入受体鼠的子宫内。受体鼠事先用激素处理使其子宫处于接受状态。分娩后，可以用 DNA 印迹或 PCR 方法从少量 DNA(最方便的是从尾部取血)中检测该动物中是否有注入的 DNA。还可以用 Northern 印迹来检测该序列的转录情况。在本例中，动物 A 和 C 已掺入了注射的 DNA，而动物 B 和 D 是阴性。C 的 DNA 拷贝数比 A 多(通常有 1～200 个拷贝)，同时也表达该基因，而 A 则不表达。转基因通常按孟德尔方式遗传给子代。图右侧显示了胚胎干细胞的进展情况。为获得预期的 DNA 变化，用体外培养选择 ES 细胞。首先从怀孕的供体鼠获取囊胚的早期胚胎，然后将上述 ES 细胞注入囊胚中。注射后的囊胚被重新植入受体鼠中。子代鼠是嵌合体，其中的一些细胞来自原来的囊胚，另一些来源于 ES 细胞。当这个嵌合体鼠与另一个鼠交配时，它会要么提供源于 ES 的、要么源于非 ES 细胞的生殖细胞。源于 ES 细胞的小鼠将带有与原始培养的 ES 细胞相同的遗传变化，因此能经交配成功地建立一个稳定的基因工程的小鼠品系。DNA 印迹或 PCR 能检测出是否存在改变后的基因。

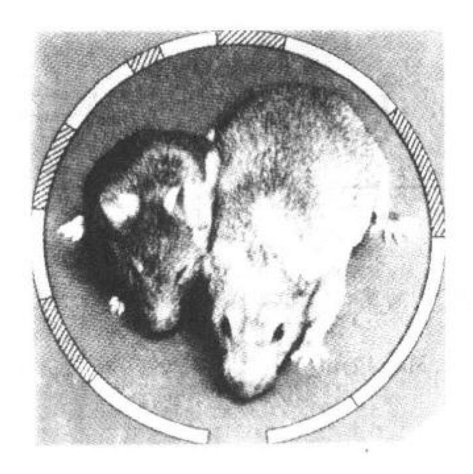


图 5-28　小鼠和超级小鼠。左侧较小的鼠是一个正常的小鼠(非转基因小鼠)；而较大的鼠是转移了大鼠生长激素基因的超级小鼠。生长激素基因在这里构建成过量表达状态。该转基因鼠的重量是正常小鼠的 2～3 倍(经 Palmiter RE, Nostedt G, Gelinas RE, Hammer RE, Brinster RL 等的许可引自 Metallothionein-human GH fusion genes stimulate growth of mice. Science 1983; 222:809～814.)。

利用这些技术，现在能生产实际上带有任何意愿改变基因组的、甚至包括将单点突变导入到特定基因靶位的转基因小鼠。利用这种方法，在已知致病基因的情况下，我们能为多种人类疾病构建强效的动物模型。这些动物可以用来研究人类疾病的发病机制以及试验新的治疗方法。

四、总　结

我们总结了基因和基因家族总体和详细的结构以及重组 DNA 技术的强有力的方法。这些技术使我们对人类分子生物学有了如此卓越的了解。这些正常结构和功能知识使我们现在易于深入分析人类疾病的分子基础，这将在下面两章中进行讨论。

推荐读物

Beardet AL, Scriver CR, Sly WS, Valle D. Genetics, biochemistry, and molecular basis of variant human phenotypes. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic basis of inherited disease. 7th ed. New York: McGraw-Hill, 1995:53～118. *A general overview of recombinant DNA technology.*

Bronson SK, Smithies O. Altering mice by homologous recombination using embryonic stem cells. J Biol Chem 1994; 269:27115～27158. *A discussion of the techniques for generating knockout mice by gene targeting.*

Grompe M. The rapid detection of unknown mutations in nucleic acids. Nat Genet 1993; 5:111～116. *A summary of methods of screening for mutations.*

Lewin B. Genes V. New York: Oxford University Press, 1994. *An excellent textbook of basic molecular biology.*

Tjian R. Molecular machines that control genes. Sci Am 1995; 272:54～61. *A revies of transcription factors and the regulation of gene expression.*

第六章 人类疾病的分子遗传学——血红蛋白病

创造发现是一件如此美妙的事物，既像坠入爱河又像登上高山之巅。

——诺贝尔奖得主 Max Perutz 在称为血红蛋白结构的高山顶峰幸遇丘比特爱神

上一章的目的是介绍人类基因的基本结构及用于分析这些基因的各种分子生物学和重组 DNA 技术。本章力求将这些知识更加集中于一个研究得很透彻的体系——即人的血红蛋白基因上，显示这些原理如何直接导致遗传病的发生。实际上，如在第一章和第五章所指出的，我们对于基因结构，它们的发育及组织特异性调控的理解的许多主要事实都是以血红蛋白基因为模式系统。因此，对于此体系来说，基础遗传学和临床医学间的通路确实是双向的说法是正确的。

一、血 红 蛋 白

红细胞是由成年人骨髓产生的，无核，它在从骨髓中释放出来之前细胞核就丢失了。年轻的红细胞被称为网织红细胞。它仍含有一些信使 RNA，但几天后就消失了。在它们以后近 120 天的生命周期中，红细胞是以膜包绕着蛋白质的弹性小袋，在此系统中循环，大约 70% 的蛋白质是血红蛋白。血红蛋白是由 2 条 α 链和 2 条 β 链组成的四聚体，由氢键自发地集合成稳定的结构（图 6-1）。每条珠蛋白链还含有 1 个血红素分子，它是涉及多个步骤的复杂酶学反应的产物而形成的一个四环状的有机复合物。血红素分子和铁一起与分子氧相结合。

血红蛋白系统所以研究得透彻是有某些原因。其一是红细胞提供了可移动和可更新的便利的物质来源。血红蛋白已以如此高浓度存在，因此不必进一步进行血红蛋白的纯化。这种同样的天然富集当然可以用于研究血红蛋白的分子遗传学；网织红细胞含有 α 和 β 珠蛋白 mRNA 是它们主要的多聚体腺苷酸化 RNA 的组分。因此如第五章所指出的，可以用一种相对直接的方式进行 α 和 β 珠蛋白 cDNA 的克隆。最后，血红蛋白研究得特别透彻是因为涉及此系统中存在大量遗传病，包括镰状细胞贫血，它是第一个“分子病”，将在下面详细讨论。

二、血红蛋白基因

人类 α 和 β 珠蛋白基因区域的模式图如图 5-2 所示。α 和 β 两个基因实际上都是基

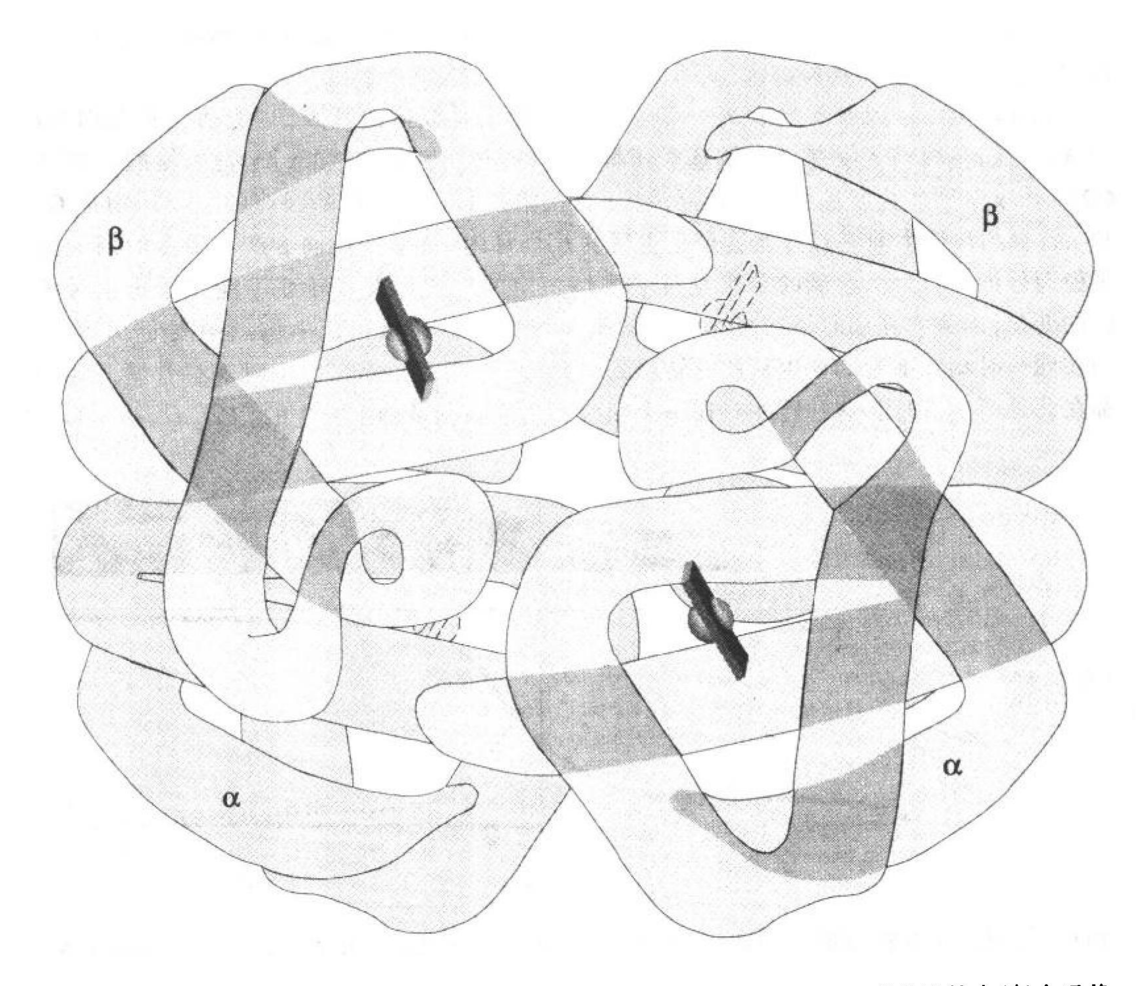


图 6-1 血红蛋白四聚体示意图。图示 2 条 α 和 2 条 β 链，每条链与一个血红素分子结合(每个亚基中央的平面结构)(引自 Dickerson RE, Geis I. Hemoglobin: structure, function, evolution, and pathology. Menlo Park, CA: Benjamin/Cummings, 1983.c W Irving Geis.)。

因簇的一部分，β 珠蛋白基因簇位于 11 号染色体上，而 α 珠蛋白基因簇位于 16 号染色体上。这些基因簇，除包括主要的成年基因——α 和 β 外，还有其他表达序列，它们在发育的不同阶段发挥作用。两个基因簇中，基因的排列按 5′至 3′转录方向与所有基因相同，而用于发育最早期的基因位于基因簇的 5′端，用于晚期的位于 3′端。正如第五章所指出的，两个基因簇都含有假基因。假基因是曾经有功能的基因残留物，经过突变使它们不能再产生蛋白质。在两个基因簇中，大部分 DNA 都位于功能未知和可能非必需的基因间区域中。

图 6-2 表示这两个基因簇中功能基因的表达模式，是按发育年龄的功能划分的。在生命的前几个星期，血红蛋白的合成发生在卵黄囊中，那时主要的血红蛋白是由 2 个 ζ 链(在 α 基因簇中编码)和两个由 β 基因簇转录而来的 ε 链组成四聚体。然而很快，这两个珠蛋白链的产物迅速消失且 α 珠蛋白基因簇开始转录 α 作为它的主要产物。实际上有 2

个 α 基因，α2 和 α1，但它们在编码区域是一致的，因此产生单一蛋白质，此后 α 基因在个体的一生中一直保留下来。

而 β 珠蛋白基因簇却含有另一套基因——胎儿血红蛋白基因，它包括一对基因 Gγ 和 Aγ，它们在编码区仅有一个氨基酸不同，在 136 位 Gγ 是甘氨酸，Aγ 是丙氨酸。如图 6-2所示，γ 基因作为胚胎的基因将要消失，而在胎儿期 β 珠蛋白基因簇的主要产物是 Gγ 和 Aγ，大约从出生前一点开始直至出生后的几个月内，从 γ 产物向 β 产物转变的调控是平稳而仔细的，因此，它俩的总量是几乎维持恒定的。这种发育转变不能归于出生时环境的变化，因为它在那时早已发生了。同样不能归于血红蛋白产生位点的改变，尽管大约在同一时刻红细胞的产生从胎肝转移到骨髓(图 6-2)。如果在出生时观察骨髓，γ 和 β 都在此处产生，因此，这种转变是在一个细胞一个细胞的基础上发生的。

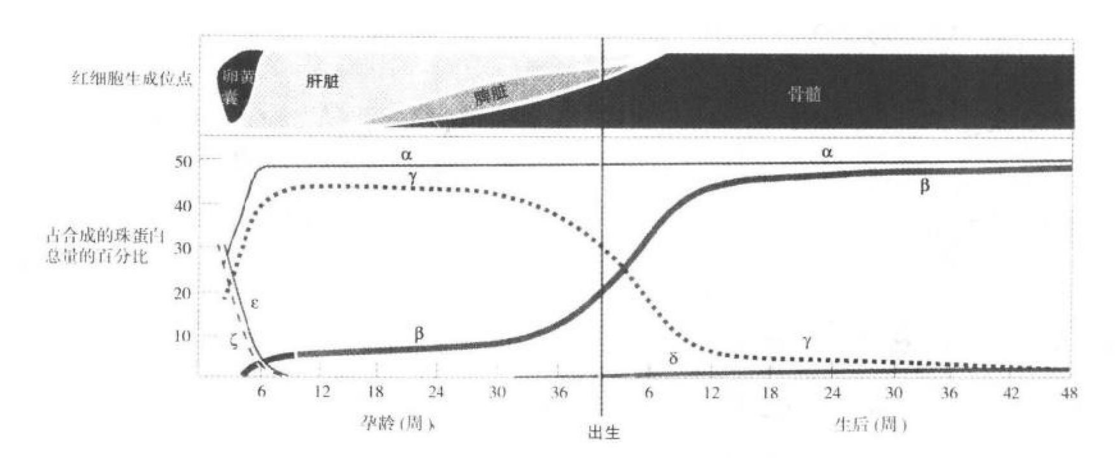


图 6-2　人血红蛋白表达的发育模式。图示产生各种珠蛋白链的时段。图上部示各阶段红细胞生成的基本位点(引自 Weatherall DJ, Clegg JB. The thalassemia syndromes. 3rd ed. Oxford: Blackwell Scientific Publications, 1981.)。

图 6-3 所示在此发育模式过程中可能存在的血红蛋白四聚体。在了解它们的分子基础之前，所有这些均已鉴定而且给予有些错误的命名。在胚胎期，主要的血红蛋白是血红蛋白 Gower Ⅰ。当胚胎基因消失而胎儿基因出现时，血红蛋白 Gower Ⅱ 和 Gower Porland 暂时出现。胎儿期，主要的血红蛋白是血红蛋白 F，它包括 $\alpha_2G\gamma_2$ 和 $\alpha_2A\gamma_2$。在成年时期，主要的珠蛋白是血红蛋白 A，它是由 2 个 α 链和 2 个 β 链构成的四聚体。还存在少量的成年人珠蛋白，称为血红蛋白 A_2，它是由 2 个 α 链和 2 个 δ 链构成的四聚体。δ 基因的发育时间和 β 相似，但它在其启动子上获得了若干改变，特别是在 CCAAT 盒区域，使之相当无效。至少部分由于这个原因，δ mRNA 的量比 β mRNA 的量低得多，最后的结果是正常成人个体有近 97.5%的血红蛋白 A，大约 2%的血红蛋白 A_2 及约 0.5%的血红蛋白 F。

γ 基因在胎儿发育阶段的表达是相对现代进化的事件，仅在灵长类动物中出现。据推测这种另外的血红蛋白有优势，因为这些物种有长的妊娠期，在这么长的时期中，需要一种更有效的方式来保证足够的氧运送到生长中的胎儿上，胎儿血红蛋白比成年人血红蛋白的氧亲和力高一些，因此可以通过胎盘从母体循环中更有效地获取氧。

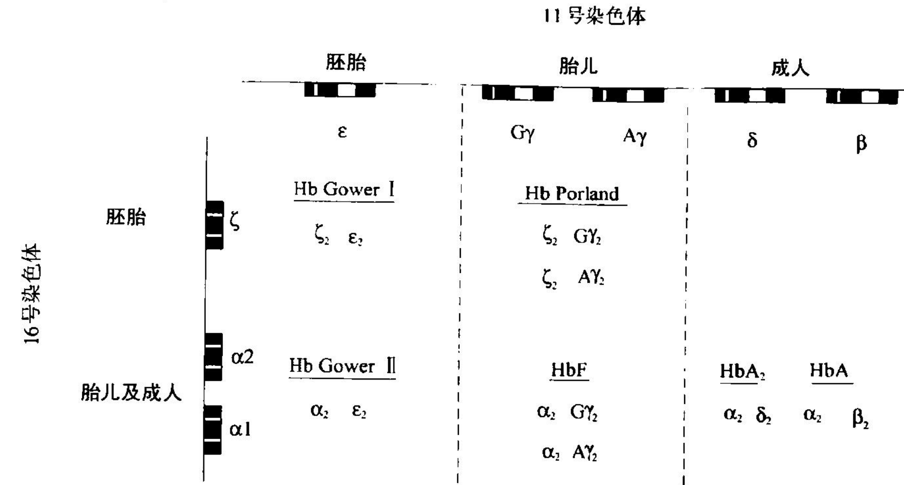


图 6-3 位于16号染色体上的α珠蛋白基因簇和位于11号染色体上的β珠蛋白基因簇可能产生的四聚体产物模式图及其历史命名。

三、β珠蛋白基因的正常结构及表达

第五章中介绍的真核生物基因的许多基本特征都首先通过研究人类β珠蛋白基因而确定,如图6-4所示。β珠蛋白基因的启动子区域包含一个TATA盒,但在这个例子中,此序列实际上是ATA。它也有一个CCAAT盒,位于它更上游的序列是CACCC,它只重复一次,而且存在于许多其他珠蛋白基因中。人类β珠蛋白基因组织特异性强,仅在红细

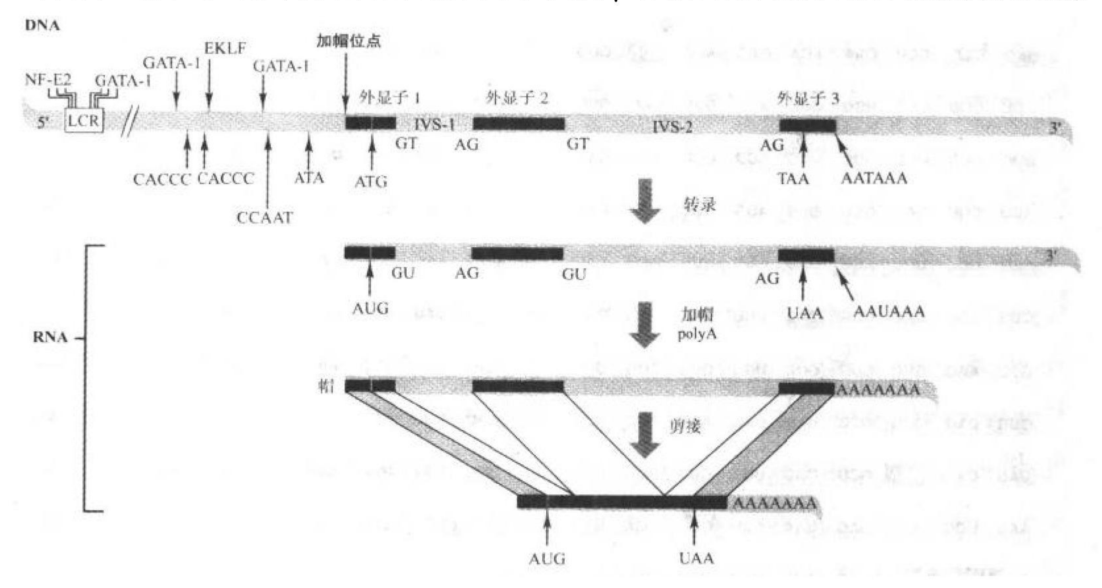


图 6-4 人类β珠蛋白基因模式图。图上部是基因组DNA模式图,图示出重要的侧翼调控序列。条框下表示启动子中重要的序列基序,条框上表示与此区域结合的几个重要的转录因子。对于β珠蛋白基因簇中所有基因的有效表达均很必要的基因座控制区域(LCR)也在图中表示出来,它位于β珠蛋白基因其本身上游的50kb以远。内含子及它们与外显子的接头,多聚腺苷酸化位点如图所示。基因转录成RNA,RNA加工成成熟的细胞质信号的过程也描述出来。

胞及其前体中才有大量表达。在过去几年中，在理解这个高度调控的表达分子基础上取得了很大进展。β珠蛋白基因启动子中的若干序列基序作为组织特异性转录因子结合的靶位点。这种因子的一个重要例子是 GATA-1。GATA 基因是转录因子的家族之一，结合到 DNA 的 (T/A)GATA(A/G)序列。在 β 珠蛋白基因启动子以及包括所有血红蛋白基因在内的许多其他红细胞特异性基因启动子中可发现许多拷贝的 GATA-1 结合基序。在胚胎干细胞中(参见第五章)，通过基因打靶方法获得的缺乏 GATA-1 的转基因小鼠不能形成血红细胞。转录因子 EKLF 结合到 CACCC 基序上，也参与控制 β 基因表达。CACCC 序列中的一个突变阻碍了 EKLF 转录因子的结合，导致了 β 地中海贫血(见下)。多个其他转录因子和 β 珠蛋白基因启动子的复杂相互作用对正确表达也是必需的。最近发现另一个重要的调控子，它与 β 珠蛋白基因相距远一些，因为它对于调控 β 珠蛋白基因簇内的所有基因都很重要，此区域被称为**基因座控制区域**(locus control region，LCR)。LCR 位于 ε 珠蛋白基因上游 6～20kb 处(参见图 5-2 及 6-4)，距离珠蛋白基因约 60kb。LCR 包括 GATA-1 和 NF-E2 在内的一些红细胞特异性转录因子的结合基序。因此 LCR 和 β 基因簇中(图 6-3)每种珠蛋白基因的启动子之间的相互作用的复杂方式，精确地调控从胚胎到胎儿最终到成人血红蛋白的转变，此复杂方式仍需进一步深入研究。正如我们将看到的，对此发育过程的操作可以给某些珠蛋白基因遗传病的治疗提供一个重要方法。

β 珠蛋白基因包括 3 个外显子和 2 个内含子，内含子 2 远比内含子 1 大。如图 6-4 所示，它有短的 5′和 3′非翻译区，加帽，多聚腺苷酸化，并以正常方式进行剪接全部过程。而最后产生的 mRNA 转运到细胞质中，它的全部序列如图 6-5 所示。值得注意的是在加

−15　−10
m^7GpppAC AUU UGC UUC UGA CAC AAC UGU GUU CAC

−5　0　5　10
UAG CAA CCU CAA ACA GAC ACC **AUG** GUG CAC CUG ACU CCU GAG GAG AAG UCU GCC GUU

15　20　25　30
ACU GCC CUG UGG GGC AAG GUG AAC GUG GAU GAA GUU GGU GGU GAG GCC CUG GGC AGG

35　40　45
CUG CUG GUG GUC UAC CCU UGG ACC CAG AGG UUC UUU GAG UCC UUU GGG GAU CUG UCC

50　55　60　65
ACU CCU GAU GCU GUU AUG GGC AAC CCU AAG GUG AAG GCU CAU GGC AAG AAA GUG CUC

70　75　80　85
GGU GCC UUU AGU GAU GGC CUG GCU CAC CUG GAC AAC CUC AAG GGC ACC UUU GCC ACA

90　95　100　105
CUG AGU GAG CUG CAC UGU GAC AAG CUG CAC GUG GAU CCU GAG AAC UUC AGG CUC CUG

110　115　120　125
GGC AAC GUG CUG GUC UGU GUG CUG GCC CAU CAC UUU GGC AAA GAA UUC ACC CCA CCA

130　135　140
GUG CAG GCU GCC UAU CAG AAA GUG GUG GCU GGU GUG GCU AAU GCC CUG GCC CAC AAG

145　150　155　160
UAU CAC **UAA** GCU CGC UUU CUU GCU GUC CAA UUU CUA UUA AAG GUU CCU UUG UUC CCU

165　170　175　180
AAG UCC AAC UAC UAA ACU GGG GGA UAU UAU GAA GGG CCU UGA GCA UCU GGA UUC UGC

CU**A AUA AA**A AAC AUU UAU UUU CAU UGC POLY A

图 6-5　人类 β 珠蛋白信使 RNA 的核苷酸序列。起始密码子 AUG 被框起来，用箭头标记上，排序为 0。终止密码子 UAA(147 密码子)也被标记上。信使的 3′端附近的多聚腺苷酸化信号也被框上，它后面的 19 个核苷酸后面接有一串 A 残基。

帽部位存在 7－甲基鸟苷酸，它后面是一小段非翻译 RNA。第一个 AUG 密码子被框上，标志着这个 RNA 翻译的开始。这个起始的蛋氨酸随后从多肽链上被切割下来，因此成熟的蛋白质的第一个氨基酸是由 GUG 编码的缬氨酸。成熟的多肽一共有 46 个氨基酸，当遇到 UAA 终止密码子时翻译才结束。RNA 上剩余的 3′部分代表非翻译序列。AAUAAA 序列在核中被用作多聚腺苷酸化的信号，它被框上，并在 3′端一边 19 个核苷酸之后连有一串腺苷酸残基。

现在我们准备看一下在珠蛋白基因中存在的大量突变。这些突变可被分为引起质量畸变（如镰状细胞贫血）及引起数量畸变（地中海贫血）的突变，这些疾病统称为**血红蛋白病**（hemoglobinopathies）。

四、引起珠蛋白质变的突变

（一）错义突变

如第二章所介绍的，错义突变是由于单个核苷酸的改变导致由特定 3 碱基密码子编码的氨基酸的改变。由于对 α 和 β 基因进行认真仔细的研究，发现了大量的错义突变，其中许多没有引起任何表型异常。如图 6-6 所示，α 和 β 珠蛋白基因的大部分氨基酸都存在已被证实的错义突变。这些突变的表型和在一些突变中所见的功能异常的关系使我们对于决定它特性的血红蛋白四聚体功能区域的理解更加透彻。对一个分子特定位点突变的结果与分子功能特性的相关性的一般研究策略具有深远的意义。

β 珠蛋白基因中最常见的错义突变在纯合子时导致镰状红细胞贫血。在黑人群体中这种疾病的发病率约为出生者 1/500，这种疾病引起严重的发病率和致死率。此病是隐性的，但携带者容易被检测出来。如第四章所介绍的，运用 Hardy-Weinberg 定律可以推算，如果此病的发病率 $q^2=1/500$，携带者频率 $2pq$ 约为 8%，而对携带者的筛查证实了这一结果。考虑到纯合镰状细胞贫血能导致寿命缩短，生育能力降低，在黑人群体中此基因达到如此高频率是惊人的。然而正如第四章所叙述的，对它的发生有明确的解释，特别是镰状细胞贫血杂合子能适当地防护由于疟原虫感染产生最严重及致死性的后果，显然这就使此基因在长期被选择。因此，即使纯合子（1/500）被选择出来，对杂合子（1/12）的适度选择已使该基因在世界的疟疾流行地区达到高频率。

镰状细胞贫血是第一个成功地在分子水平得以解释的人类疾病。镰状细胞贫血的基本病变是 Linus Pauling 在 1949 年首先推论出来的。通过对血红蛋白进行仔细分析，Vernon Ingram 随后发现镰状细胞 β 珠蛋白基因在第 6 位点处有一单个氨基酸的改变，此位点正常的残基应为谷氨酸，而在镰状细胞贫血中为缬氨酸替代。随后，对各基因的克隆和测序揭示了预期的 DNA 序列改变，即第 6 密码子的序列从正常基因的 GAG 改变为镰状细胞贫血的 GTG。有意义的是，另一种相对常见的被称为血红蛋白 C 的 β 珠蛋白链错义突变也是由第 6 密码子改变所致，由 GAG 变为 AAG，使此位点成为赖氨酸（图 6-7）。

确实值得注意的是人类单倍体基因组中 3×10^9 个核苷酸中一个单核苷酸的改变就可使镰状细胞贫血患者产生一系列严重的临床问题，这些问题的基础如图 6-8 所示。根

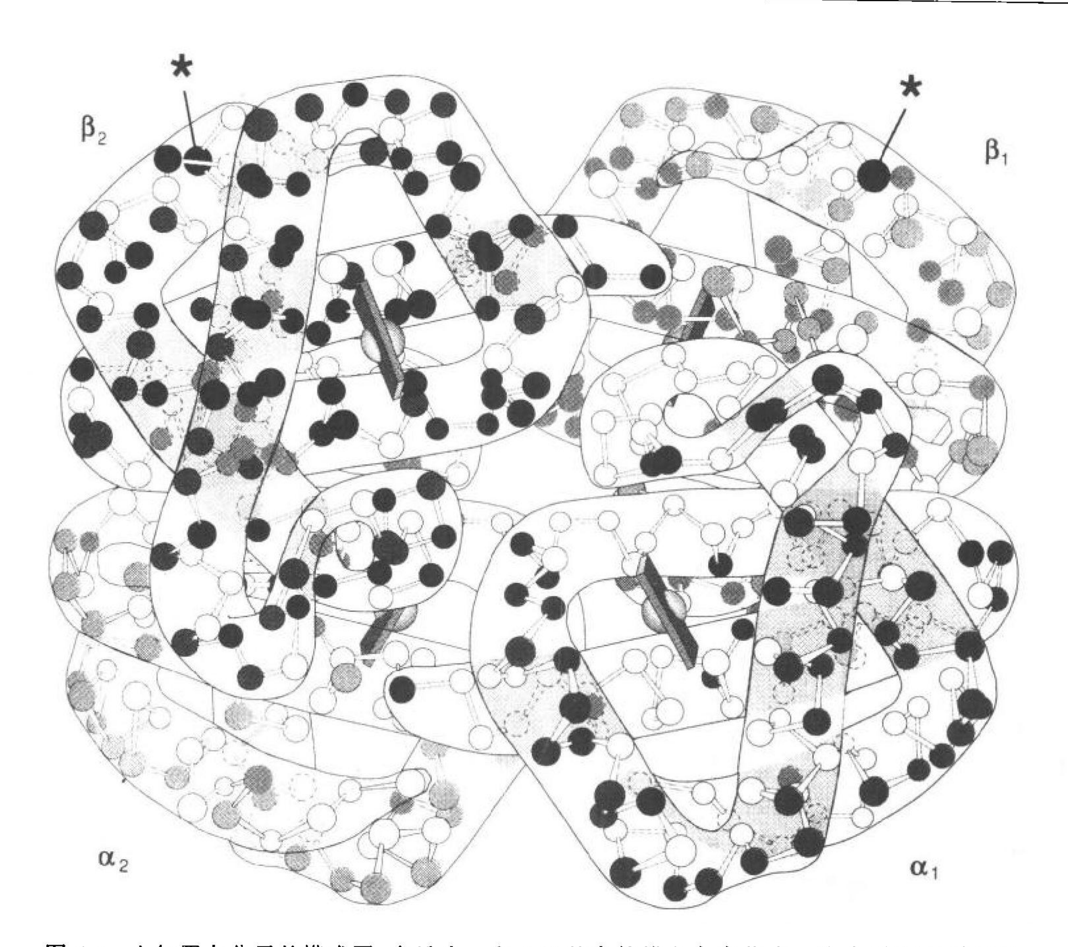


图 6-6 血红蛋白分子的模式图，表示出 α 和 β 亚基中的错义突变位点。这个图基于到 1982 年为止所报道的所有血红蛋白突变，那时，已发现大量的异常血红蛋白，包括 α 和 β 链中总共 287 个氨基酸位点中的 170 个。阴影圈示发现氨基酸改变的位点，在 α_1 和 β_2 这两个最靠近的亚基中，导致珠蛋白功能异常的突变用黑色显示。镰状细胞贫血中 β_6 位点用星号表示。到 1996 年为止，共鉴定出 586 种 α 和 β 链变异体，包括 287 个氨基酸位点上的 238 种(引自 Dickerson RE, Geis I. Hemoglobin: structure, function, evolution, and pathology. Menlo Park, CA: Benjamin/Cummings, 1983.)。

本原因是缬氨酸替代谷氨酸导致血红蛋白四聚体在脱氧时，聚集成阵列而不是保持可溶解状态。如图 6-8 所示，这些阵列在结构上几乎可变为结晶体，因此，每当一个携带镰状血红蛋白的红细胞通过循环并变为脱氧状态，其中的血红蛋白聚集，随后使红细胞变形，并相对无弹性，而且不能通过直径比红细胞本身小的纤细的毛细血管床。经过脱氧及再携氧的重复循环，一些细胞变成不可恢变性的镰状，结果阻塞了微循环，特别是在氧压力低的地方。骨严重受累，在镰状细胞“危象”中可导致频发的严重骨痛。长此下去，这种频繁的阻塞微循环导致内脏器官的严重损伤，特别是心脏、肺和肾脏。脑血管意外是另一种严重的并发症，因为破坏的红细胞包含血红蛋白结晶，镰状细胞贫血患者常有慢性贫血，就更加重了心脏负荷。

虽然我们对于镰状细胞贫血的分子机制的理解取得显著进步，但治疗上的进展令人

失望。在提高对这种很常见且危害生命疾病的治疗方面，进展如此缓慢为现代分子医学敲响了警钟，即现代分子医学中的基础研究的迅速进展很难直接转化为临床效益。尽管几十年来治疗镰状细胞贫血患者有所挫折，最近的研究结果给予我们一些希望。如最近选择严重的患者采用一种温和的化疗法(羟基脲)，可使胎儿血红蛋白产量增加并改善红细胞功能。此外，现在一些病人可以用骨髓移植的方法治愈。此方法伴有严重的风险，所以它的使用目前仅限于最严重的患者。

DNA			
	密码子 5	6	7
β^A	...CCT	GAG	GAG...
β^S	...CCT	GTG	GAG...
β^C	...CCT	AAG	GAG...

蛋白质			
	5	6	7
β^A	...Pro	Glu	Glu...
β^S	...Pro	Val	Glu...
β^C	...Pro	Lys	Glu...

图 6-7 导致镰状血红蛋白(β^S)和血红蛋白 C(β^C)的密码子 6 上的 DNA 及蛋白质异常。

当存在典型的临床症状时，通常采用检测血涂片法很容易对镰状细胞贫血诊断(图 6-8)。通过将它们的红细胞暴露在氧压力极低的地方可以诊断出携带者，在这种情况下，即使杂合子来源的血细胞也可成为镰状。重要的是，杂合子(常指具有镰状特征)无症状，除非可能像高空无压力飞行时这样的极低的氧压力。

然而，为确定诊断常进行血红蛋白电泳。如图 6-9 所示，电荷不同的血红蛋白可通过在电场中泳动而分离，因为镰状突变将负电荷氨基酸(谷氨酸)变为中性(缬氨酸)，结果在电场中蛋白质在凝胶上的泳动就有明显改变。这样的凝胶也可检测血液中少量的血红蛋白 F 和血红蛋白 A_2，它们的位置和血红蛋白 A 不同。如上所示，血红蛋白 C 与血红蛋白 A 相比有 2 个电荷差异，因为在第 6 密码子处由负电荷氨基酸变为正电荷氨基酸(赖氨酸)，导致在凝胶上向正极泳动的条带比血红蛋白 S 更慢，且距离血红蛋白 A_2 的位置更近。图 6-6 所示的许多其他错义突变最初就是通过血红蛋白电泳中存在异常条带而被检测。

如有可用的红细胞时，血红蛋白本身的电泳是诊断这些异常的最直接方法。然而，在有些情况下，如产前诊断，从被测者得到纯血液样品是有风险或不可能。因为编码 β 珠蛋白基因的 DNA 存在于有机体的所有细胞中，因此通过直接分析来自任何可用的细胞 DNA 是可能诊断镰状突变的存在。如在以后章节中介绍的，可以通过羊水或绒毛活检得到这种细胞，远比获得胎儿血样要容易得多。

图 6-10 所示在 DNA 样品中检测镰状突变的一种方法。限制性酶 *Mst* Ⅱ 识别序列为 CCTNAGG，N 代表任意核苷酸。因此密码子 5，6 及密码子 7 的第一个碱基序列通常存在一个此酶的识别位点。然而，在镰状突变时，使此 *Mst* Ⅱ 位点消失，因为此突变是 A 变为 T，使此区域不再匹配这个限制性内切酶。这种情况是有些偶然。限制性内切酶识别位点并不存在于所有的位置，只有一些单个碱基的改变引起人类疾病改变了限制性位点才可用此方法直接检测。由于镰状突变破坏了限制性位点，可以用 *Mst* Ⅱ 切割它们的 DNA，并用 DNA 印迹杂交分析结果才能鉴别杂合子与纯合子，如图 6-10 所示。正常 β 珠蛋白的等位基因在密码子 6 上存在 *Mst* Ⅱ 位点，当与来自此基因 5′侧翼区域序列的探针杂交时出现 1.15kb 带。然而，镰状等位基因的第 6 密码子识别位点消失，相邻的下一个位点位于 3′端方向以远 200bp 处(在 β 珠蛋白基因内)，因此此片段为 1.35kb。正常纯合子个体，

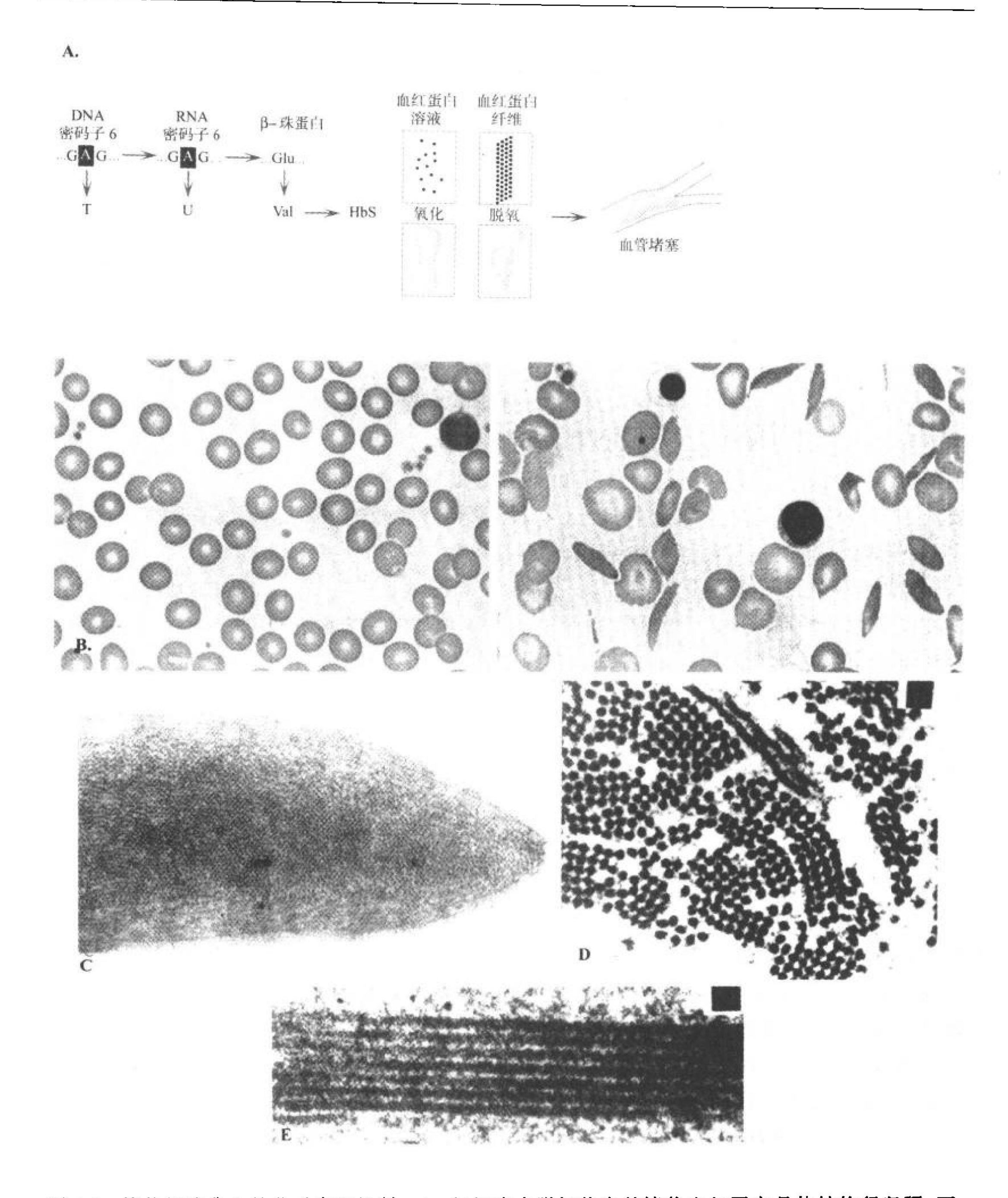


图 6-8 镰状细胞贫血的分子病理机制。A. 红细胞内脱氧状态的镰状血红蛋白晶体结构很坚硬，不能通过小的毛细管。B. 右侧血涂片中可见镰状细胞，左侧血涂片为正常对照。C. 在横向(D)和纵向(E)切片上，在脱氧镰状细胞电镜图中可见血红蛋白纤维(引自 Bunn HF, Forget B. Hemoglobin: molecular genetic, and clinical aspects, Philadephia: WB Saunders, 1986.)。

杂合子及纯合受累个体可用此方法区分出来。实际上，这种分析方法最常用 PCR 技术来进行。如第五章所叙述的，用与突变的密码子 6 的两侧相匹配的一对 PCR 引物扩增 DNA

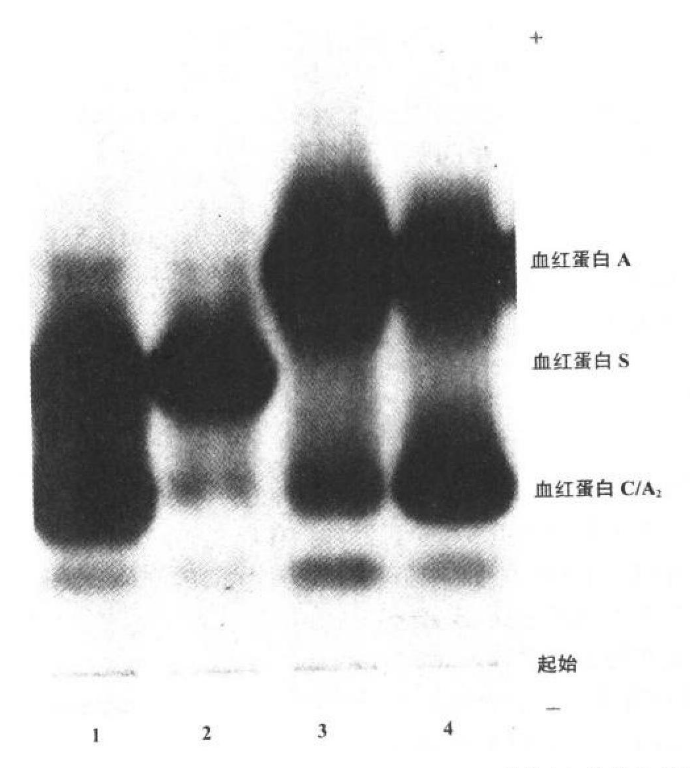


图 6-9　血红蛋白电泳。血红蛋白样品加在标记为起始处的位置,然后放置在电场中,泳动的距离反映了分子的总的电荷。这就能区分血红蛋白 A,S,C 及 A_2(C 和 A_2 泳动得十分靠近)。正常个体(AA,带 3)主要含血红蛋白 A 及一小部分的血红蛋白 A_2,镰状纯合子(SS,带 3)只有血红蛋白 S 和 A_2。图示 AC(带 4)及 SC(带 1)个体的血红蛋白样品(引自 Weatherall DJ,Clegg JB. The thalassemia syndromes. 3rd ed. Oxford: Blackwell Scientific Publications, 1981.)。

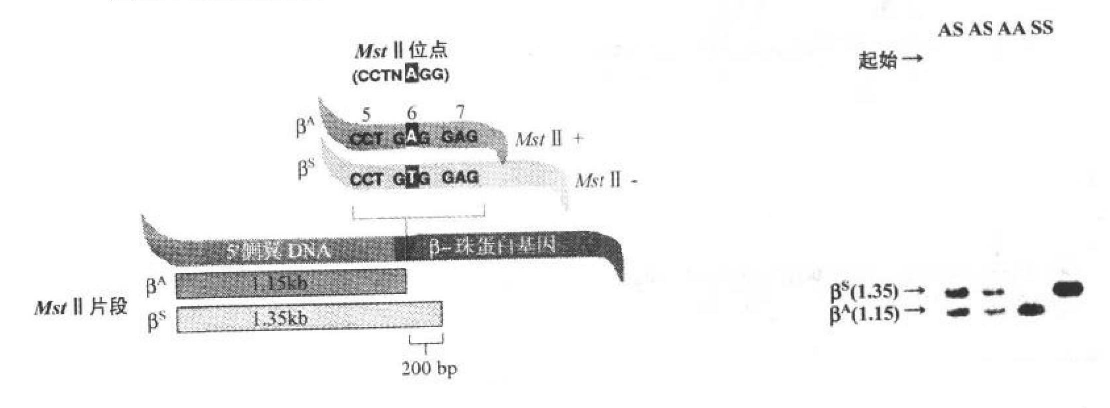


图 6-10　用 DNA 印迹分析诊断镰状突变。图示在 5,6 和 7 密码子由于镰状突变使正常 *Mst* Ⅱ位点消失。右侧是纯合子 A(AA),纯合子 S(SS)及 AS 杂合子个体 DNA 的 NDA 印迹杂交结果。探针是来自 5′侧翼区 DNA 的基因组片段(引自 Bunn HF, Forget B Hemoglobin: molecular genetic, and clinical aspects, Philadelphia: WB Saunders, 1986)。

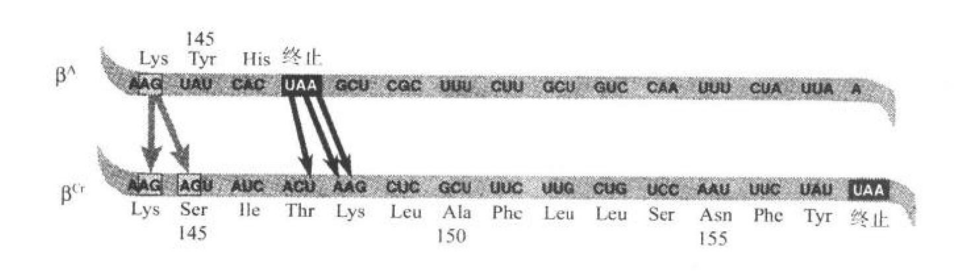
片段，然后用 *Mst* Ⅱ 消化，直接在凝胶上分析。总之，因为有了 PCR 也可以用许多其他方法，而不必要通过选择能区分正常及受累等位基因的限制酶。图 5-24 示用聚合酶链式反应（PCR）和等位基因特异性寡核苷酸（ASO）的例子。

镰状细胞贫血是少数几个可以用这种直接方法进行 DNA 诊断的疾病。这可能是因为(a)在 DNA 水平上已知突变；和(b) 所有患者都具有相同突变，即没有等位基因异源性。因为大部分其他遗传病，如在本章后面将要讨论的地中海贫血综合征，多个不同的突变都可产生同一疾病的表型。在这些病例中，一般必须采用第五章中介绍的筛选方法之一。这个过程相对较为复杂，由于可能会丢失一类突变或鉴定的氨基酸替换并不就是引起疾病的真正突变。

（二）移码突变

正如第二章所叙述的，当有非 3 的倍数的少量核苷酸的插入或缺失时，就产生移码突变。由这种突变基因而来的 RNA 进行翻译时，移码突变位点后的氨基酸顺序完全窜改了。这会产生几种可能的结果。如果移码突变发生在编码区的 5′端附近，此蛋白质和它的正常的产物可能没有相同之处，或者可能迅速降解且不可能予以辨认。它的最终结果是可辨认的蛋白质产物的严重缺乏，所产生的地中海贫血类型我们将在随后详细说明介绍。然而，如果移码突变位于蛋白质末端，其结果是蛋白质的质变在多肽链的羧基端得到一段全新的氨基酸。例如，图 6-11 所示的血红蛋白 Cranston（异常血红蛋白常由被发现异常的第一个病人所在的地理位置命名）。在此 β 珠蛋白 144 和 145 密码子之间有一

A.



B.

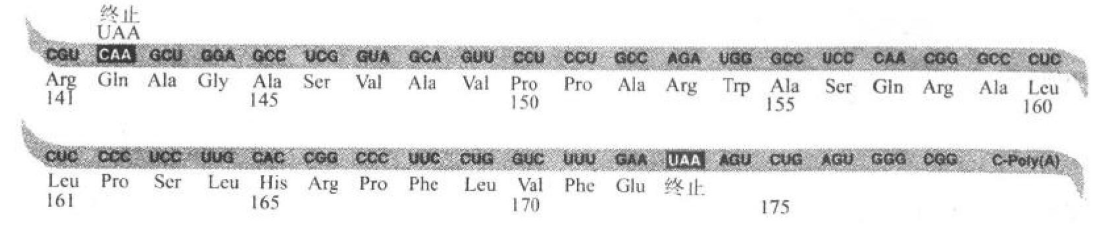


图 6-11 正常终止密码子的丢失。A. β 珠蛋白基因末端附近的移码突变导致血红蛋白 Cranston($β^{Cr}$)。B. 正常 α 珠蛋白 142 终止密码子的突变引起血红蛋白 Constant Spring，其在羧基末端有额外 31 个氨基酸。

个 2 - 核苷酸插入，它几乎在编码区的末端。此结果使正常终止密码子 6 个核苷酸远移到框外，因而不能终止阅读。因此，翻译显然进入到正常的 3'非翻译区，在偶遇下一个终止密码子之前产生了一个由 157 个氨基酸组成的蛋白质。

在血红蛋白 Constant Spring 中出现相似情况，它是 α 珠蛋白基因的一个相似突变。在这种情况下，正常终止密码子 UAA 突变为 CAA，编码谷氨酸而不是终止的密码子，因此这不是真正的移码突变而被认为是无义突变的相反类型。结果是翻译继续进行，到 α 珠蛋白 RNA 的正常 3'非翻译区内，在 173 密码子处遇到框内终止子之前继续翻译另外 31 个氨基酸。因此，所产生的 α 珠蛋白链的前 141 个氨基酸是完全正常的，但含有异常的羧基端末尾。它的出现使 α 珠蛋白有些不稳定，因此结果是质和量结合异常的 α 珠蛋白产物。

另一种有重要意义的突变来源是称为**不等交换**(unequal crossing over)的现象。如第二章所叙述的，同源染色体在减数分裂 I 期基于染色体上全长的相似序列排列起来。此时，在同源染色体之间发生交换。如果存在一段重复的核苷酸序列，就可能在排列过程中发生错误，如果染色体间发生了交换是错配的，其最终结果将是缺失或重复。图 6-12 表示了一个例子。在 β 珠蛋白基因的 91 ~ 94 密码子和 96 ~ 98 密码子中发现一个几乎完全重复的序列 GCTGCACGTG。如果此序列的第一个拷贝和第二个拷贝(而不是和它本身)排列起来，并在此区域内发生交换，其结果是氨基酸 91 ~ 95 的缺失或重复。事实上已发现此缺失并命名为血红蛋白 Gun Hill。这 15 个核苷酸的缺失使阅读框架保持完整，因此，翻译的蛋白质从氨基酸 96 以后仍是正确的。

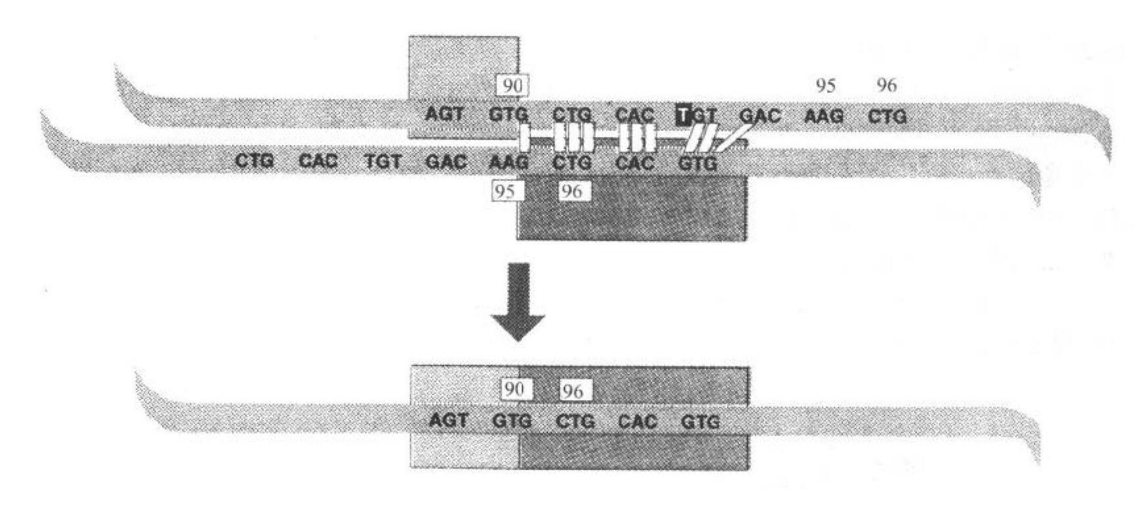


图 6-12 减数分裂时，同源区域的不等交换产生血红蛋白 Gun Hill。

δ 和 β 珠蛋白基因彼此大约 90% 同源，如果在减数分裂时它们发生错误地排列，并产生交换，则出现很大片段的缺失。图 6-13 即表示这一事件，由交换产生的染色体(用黑线表示)是由 δ 基因的 5'端和 β 基因的 3'端组成的杂种成人珠蛋白基因。已发现几种这样的变异体，依其交叉位点的精确位置而定，都被称为 Lepore 血红蛋白。因为正如上述，δ 珠蛋白基因的启动子有一些使它相对无效的突变，Lepore 血红蛋白产量下降，因此这是又一个质和量结合异常的例子。

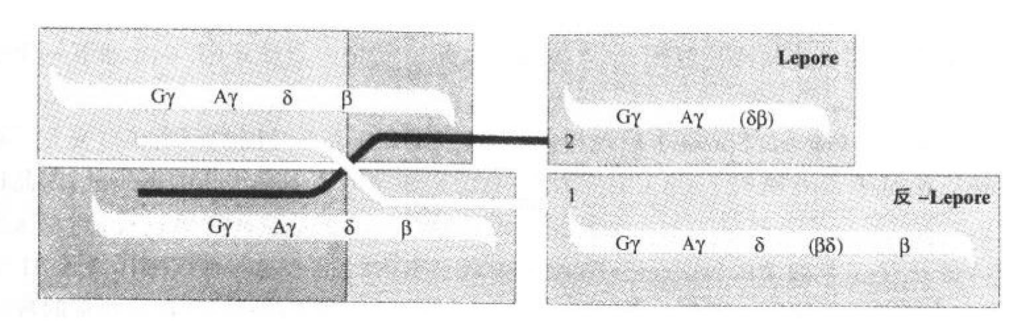


图 6-13　δ和β基因的不等交换产生血红蛋白 Lepore(杂种δβ基因)及反-Lepore(杂种βδ基因)。

值得注意的是这种不等交换还可能有另一种结果。如图 6-13 中 1 线所示，结果产生了δ和β基因完全的染色体，在它们之间插入一个杂种的βδ基因。在这种情况下，杂种基因有β基因的 5′端序列和δ基因的 3′端序列，这被称为“反 - Lepore”血红蛋白。有这种染色体的个体将产生相对正常量的δ和β，但另外还产生新的反 - Lepore 血红蛋白。

五、珠蛋白链不平衡——地中海贫血

到目前为止我们一直探讨由于这些编码区的改变而导致的δ和β珠蛋白链的主要质量异常。另一方面，地中海贫血是遗传性的血红蛋白链产量的异常，主要障碍是数量的缺乏，或由于β链则导致β地中海贫血，或由于α链则导致α地中海贫血。地中海贫血来源于希腊字“thalasa”，意为“海”。由于居住在地中海周围的居民这些病的发病率相当高，因此这个名称被沿用下来。实际上，地中海贫血不仅在地中海地区，在非洲及东南亚部分地区也是常见的，同镰状细胞贫血一样，地中海贫血的分布与疟疾的发病频率相一致，这些区域的高频率地中海贫血等位基因是在感染疟疾原虫时，反映了在这些情况下杂合子具有的一种优势。

图 6-14 表示地中海贫血的发病机制。在正常情况下，产生相同量的α和β珠蛋白链，因此它们能够在化学定量上结合形成适宜的血红蛋白四聚体。结果正常成人红细胞

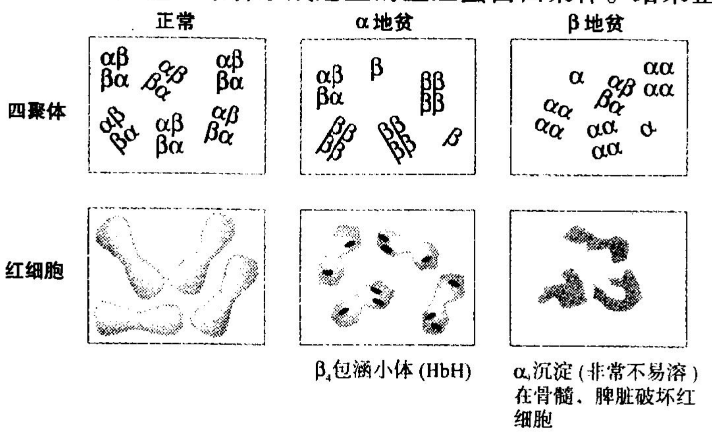


图 6-14　地中海贫血发病机理的模式图(RBCs, red blood cells)。

中含有高浓度的血红蛋白成分，平均一个细胞容积约为 $90mm^3$，α 地中海贫血的特征是 α 珠蛋白链相对缺乏而 β 珠蛋白链产量正常。如果一些 α 链仍能产生，则会形成少量正常四聚体，但将有大量过剩的 β 链。在这种情况下，β 能形成同源四聚体（β_4）。这种血红蛋白称为**血红蛋白** H(hemoglobin H HbH)，在患 α 地中海贫血的患者的红细胞中可见包涵小体。作为运氧载体，HbH 的功能显著下降。β 链的过剩及 α 链的缺乏结果使红细胞体积缩小（$50\ mm^3 \sim 80mm^3$，依病情严重程度而定），数量减少，导致贫血。

β 地中海贫血缺乏的是 β 珠蛋白链。在这些情况下，α 珠蛋白链过剩，也能够形成同源四聚体。然而在红细胞中这些四聚体很难溶解且易沉淀，使它们在成熟前在骨髓中就被破坏掉，而且被脾脏大量捕获。同样，β 地中海贫血患者的红细胞体积减小（$50\ mm^3 \sim 80mm^3$），数量也减少。

（一）α 地中海贫血

α 地中海贫血的遗传学，起初令人困惑，因为这种疾病的患者的 α 珠蛋白量似乎波动范围很大，从完全没有（导致死产）到非常接近正常水平。直到真正了解每条 16 号染色体上有 2 个功能性 α 珠蛋白基因，导致正常情况下共有 4 个 α 基因时，才将大量的混淆弄清楚。α 地中海贫血一般涉及任何从一个到全部四个基因的失活，因此，病情严重程度波动范围大也是可以理解的。图 6-15 中显示 α 珠蛋白基因失活可能出现的后果。如果四

表型	基因型	
	图解	简写
正常	α α α α	αα/αα
杂合性 α 地中海贫血 2 “静止型携带者”	α ▭ α α	α–/αα
杂合性 α 地中海贫血 1 “α 地贫特征”	▭ ▭ α α	– –/αα
黑人中 α 地中海贫血 1 表型 纯合性 α 地中海贫血 2 “α 地贫特征”	α ▭ α ▭	α–/α–
HbH 病 ($HbH=\beta_4$)	α ▭ ▭ ▭	α–/– –
胎儿水肿 血红蛋白 Bart's($=\gamma_4$)	▭ ▭ ▭ ▭	– –/– –

图 6-15 α 地中海贫血的分级及严重性，从四个基因都有功能的正常情况到无 α 珠蛋白产物的致死性的胎儿水肿。

个基因中三个有功能，临床异常极其轻微，这样的个体完全无症状，这称为“静止携带者”，它也称为α地贫2。比较明确的基因型是αα/α-或α-/αα(图6-15)。如果四个珠蛋白基因中有2个失活，这称为“α地中海贫血特征”或α地贫1。它的产生有两种可能方式，一种情况是，通常认为导致东南亚α地中海贫血，一条染色体上的2个α珠蛋白基因全部缺失，而另一条染色体上的2个基因正常(αα/—)。而患α地中海贫血1的黑人通常有不同的组合，每条染色体上各有一个α珠蛋白基因是有功能的(α-/α-)，因此这些个体是α地中海贫血2染色体的纯合子。α地中海贫血1表型相对良性，平均细胞体积下降，但这些个体相对无症状。

更严重的情况是四个基因中仅1个有功能，它的产生可能是，例如，双亲之一(α-/αα)携带一个α地中海贫血2染色体，另一个携带东南亚地区类型的α地中海贫血1(αα/—)。如果双亲将他们的地中海贫血染色体传给一个孩子，那么这个孩子(α-/—)的α珠蛋白量就会明显缺乏，最终结果是β链具有4:1的优势，导致很容易就能检测到β_4四聚体。实际上，大量的β_4的存在导致血红蛋白H病，这些患者在出生时就有中到重度的贫血，细胞平均体积非常低(大约50)，然而这并不是致死的疾病。

最严重的类型是无功能的α珠蛋白基因(—/—)，它导致死胎或早期新生儿死亡。这种情况下，主要的胎儿血红蛋白是γ链四聚体，称为**血红蛋白 Bart**(hemoglobin Bart's)。这种血红蛋白实际上没有运氧能力，因此除一少部分溶解于血液中的氧外，胎儿组织严重缺氧。此外，还有严重的贫血。心力衰竭是由于不能产生氧的心脏将溶于血中的少量不溶解的氧泵到严重缺氧的组织中的后果。因为心脏衰竭，发生显著的水肿(胎儿水肿)，而这种情况通常不能生存。只有双亲都携带一个—/染色体才会出现这种基因型。最常见的源自双亲都是东南亚地区类型的α地中海贫血1(αα/—)，当然偶尔是双亲之一携带血红蛋白H病(αα-/—)。胎儿水肿在东南亚来源的个体中最常见，在黑人群体中相对罕见，这是因为黑人群体中大部分α地中海贫血等位基因是α-/型。

α地中海贫血的病因其分子水平是多样的，最常见的异常是导致一个染色体上的一个α珠蛋白基因的缺失，其原因是染色体不等交换。如图6-16所示，α1和α2珠蛋白基因内及其周围的核苷酸都有高度同源性，图中所示X和Y区表示此区域相互之间90%以上是一致的，而Z区99%以上一致。因此这是不等交换一个理想的模式图。如果这两条同源染色体排列不正确，一个α1基因的Z区和另一个α2基因的Z区排列起来，这个间区的交换可产生两种结果，如图的上部所示，一种情况是来源于α2单个基因的5′端和来源于α1基因的3′端之间发生一段缺失。实际上，大部分α地中海贫血2等位基因都是由于此机制而产生的。而此事件的另一种情况是导致3个α基因簇集在一条染色体上。染色体的确切结构取决于交换点，在X、Y、Z区都可以发生。在大多数人群中，可见三个α基因的染色体的频率相当地高。这正好支持了这种假说。

然而，这种机制不能用来解释一条染色体上2个α基因都丢失(—/)。这样的染色体常发现有隐匿性缺失，在这种情况下，缺失很大，丢失2个α基因，这不能简单地用不等交换来解释。最后，在一些情况下，一个α基因功能的缺失是由许多轻微突变所致，如无义突变或移码突变，但同β地中海贫血(见下)比较起来，这些突变都是少数的。

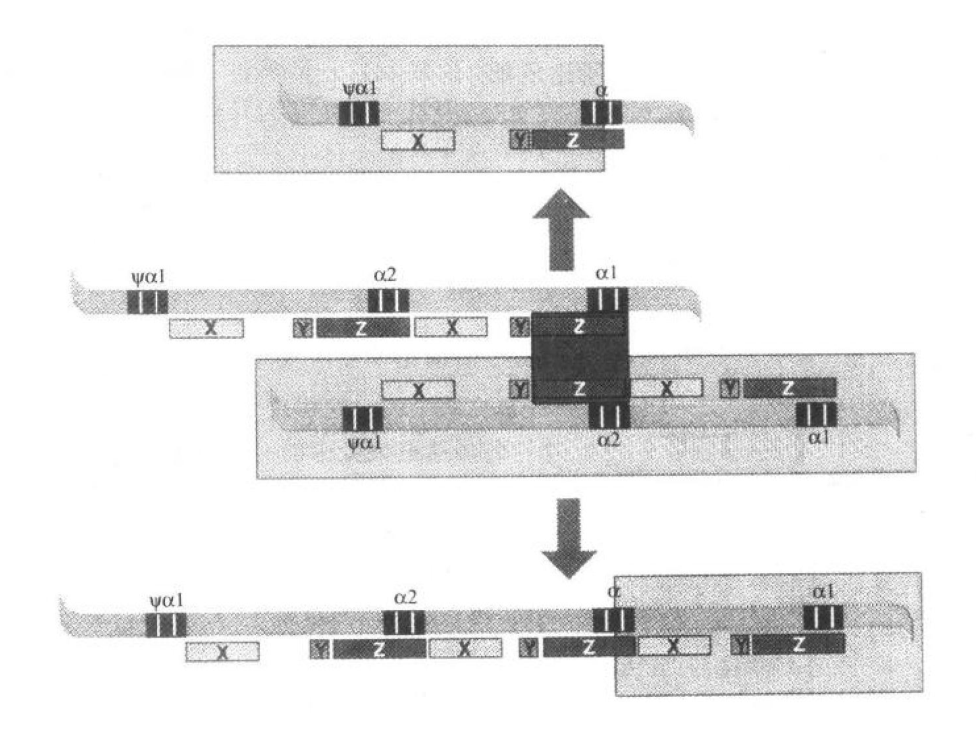


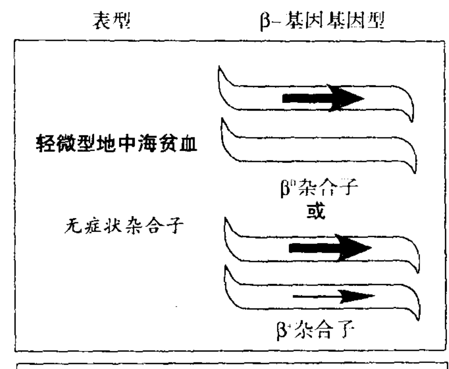
图 6-16 减数分裂时 α 珠蛋白基因簇错排导致 α 地中海贫血。图中间示不正确排列的两条染色体，一条染色体上的 α1 与同源染色体上 α2 相匹配，如果在 Z 区发生交换，结果产生仅含有 1 个 α 基因的染色体(见上)或有 3 个 α 基因的染色体(见下)。

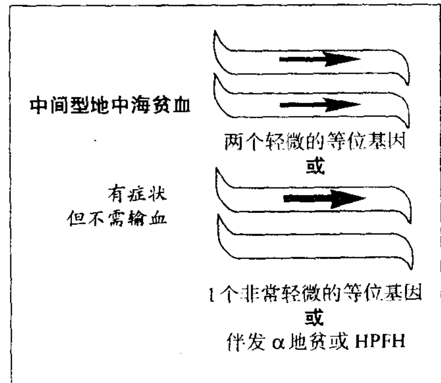
(二) β 地中海贫血

β 地中海贫血是由于 β 珠蛋白链缺乏所致。因为每条 11 号染色体上仅有 1 个 β 珠蛋白基因，不等交换的可能性就大为降低。在 δ 和 β 之间发生不等交换时，产生前面提到的血红蛋白 Lepore；因为 δ 和 β 之间的同源性很少于 2 个 α 珠蛋白基因，这是较为少见的事件。因此，发现 β 珠蛋白基因功能的缺失或下降一般是由于另外的机制引起的。

可导致 β 珠蛋白基因功能下降或缺失的大量突变使 β 地中海贫血的遗传学复杂化。此病以常染色体隐性方式遗传。然而根据红细胞体积减小，血红蛋白电泳时血红蛋白 A_2 和血红蛋白 F 具有特征性轻度增加的事实，携带者可以检测出来。如同 α 地中海贫血一样，由于存在大量可能的突变和在理解 β 地中海贫血的分子机制以前所形成的历史性术语。对其遗传学分析多少有些复杂。

图 6-17 示由 β 珠蛋白基因座突变产生的可能表型和能够解释它们的基因型。从最严重的开始，β 地中海贫血主要代表纯合的状态，即两个 β 珠蛋白基因都含有突变，使它们不能产生正常量的 β 珠蛋白。如果两个 β 基因全无功能，则都没有血红蛋白 A，这称为 $β^0$ 地中海贫血，如果有 1 个或 2 个突变，这样仍能产生少量的 β 珠蛋白，那么就称为 $β^+$ 地中海贫血。严重型 β 地中海贫血患者依赖输血，且遇到许多主要医疗问题，将在下面作介绍。





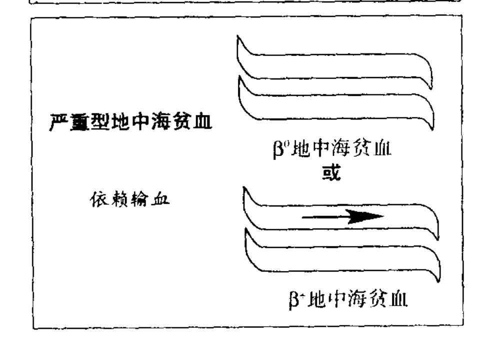


图6-17　产生严重型β地中海贫血，轻微型β地中海贫血和β地中海贫血中间型表型的可能基因型。加粗的箭头表示每一β珠蛋白基因转录的水平。值得注意的是同时伴发α地中海贫血可以降低病情严重性，其原因是改善了链合成的平衡（图6-14）。遗传性胎儿血红蛋白持续性增多症（HPFH）由于增加的γ链可以补偿缺欠的β，对β地中海贫血有益。

β地中海贫血杂合子认为是"轻微的地中海贫血"，无症状。一般来讲，他们携带一正常β珠蛋白基因及一突变基因导致其功能下降或破坏。通常在这样的个体中不能通过判断β珠蛋白水平说明是否携带地中海贫血等位基因β^+还是β^0突变，因为在这两种情况下正常的染色体都产生大部分的β珠蛋白。

有些容易混淆的术语是"地中海贫血中间型"。它曾用于指临床上有严重贫血症状但不需要输血的个体。这样患者的两个β珠蛋白基因均异常，但这些突变中的一个或2个比较轻微，因此仍能产生大量的β珠蛋白。中间型地中海贫血的另一种机制是个体具有α地中海贫血并伴有β^+地中海贫血。从图6-14所示是可鉴定的，如果两个α和β珠蛋白产量都下降，其表型较轻微，因为在红细胞中缺乏破坏性α_4沉淀物。

β地中海贫血通常是由于许多轻微突变的结果，而α地中海贫血主要是由于缺失引起的。过去十年认真的工作证实了能产生这种表型的一系列突变事件，现在已有100多种突变表明可以引起β地中海贫血，因此这种疾病是等位基因异质性的一个主要例子。由于该突变种类繁多，大部分严重型β地中海贫血的个体实际上是复合杂合子，他们从每个双亲遗传了不同的突变。

图6-18示一些β地中海贫血突变的图解。β珠蛋白基因突变遍布基因的全长，包括启动子所在的5′侧翼区域。从此分析中得到一点认识即此基因的每一部分都能是地中海贫血突变的靶点。因此这是"基因组Murphy定律"的一个例证——任何事如果可能变坏，它就一定能变坏。

没有必要了解这一系列突变事件的全部详情，但对某些类型稍加详细说明还是有用的。一个有意义的类型是个体在基因启动子处存在点突变，如图6-19所示。显示在ATA盒已发生6种不同突变，以及在CACCC序列中也有6种突变，此序列以前被认为是β珠蛋白基因的稳定部分。具有这些突变中任意一种的个体，从突变基因转录而来的β珠蛋白mRNA仅约有10%的正常量，这是这些序列在β珠蛋白基因有效转录过程中起重要作用的基本证据。而至今在CCAAT盒中尚未证实存在突变。在加帽位点（mRNA的第一个核苷酸，图6-18）

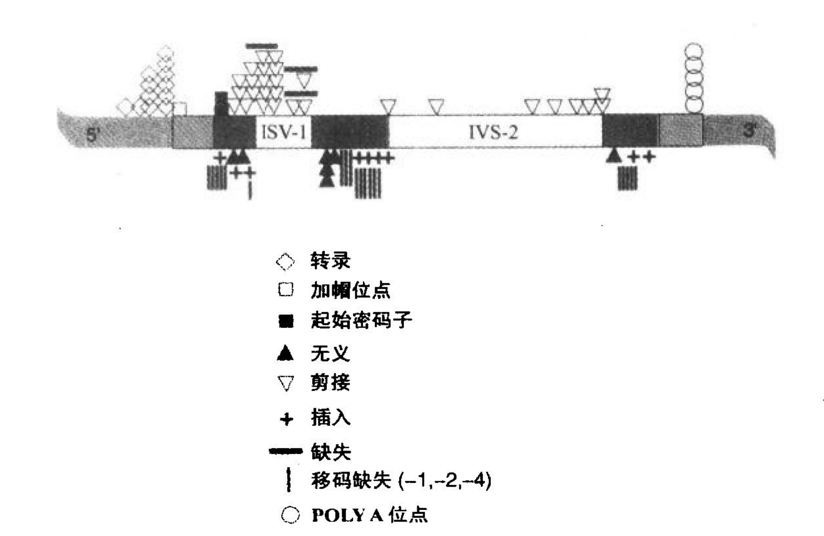


图 6-18 导致 β 地中海贫血的一些突变的模式图(引自 Kazazian HH Jr. The thalassemia syndrome: molecular basis and prenatal diagnosis in 1990. Sem Hematol 1990;27:209 ~ 228.)。

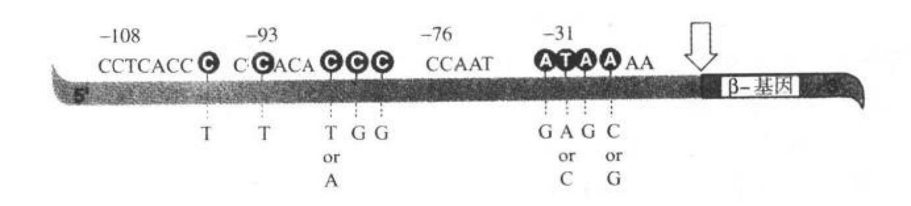


图 6-19 β 珠蛋白启动子上的突变能引起 β^+ 地中海贫血。

证实有一单个突变,这个突变可能会影响转录效率或干扰加帽过程,该过程对 mRNA 稳定性是必要的。

一些突变影响 β 珠蛋白 mRNA 有效地翻译成蛋白质,包括翻译起始密码子本身的突变(图 6-18)。链终止子突变可使翻译于成熟前终止,这可能是由无义突变或移码突变引起的结果。图 6-18 中的大部分情况是畸变发生在编码区非常前的地方,所生成蛋白质的结构丧失一半或更多,并导致整体无功能。这和上述的血红蛋白 Cranston 和血红蛋白 Constant Spring 不同,它们的畸变发生在十分靠近蛋白质末端,因此仍保留功能。在编码区前端出现的终止子突变通常产生很不稳定且在细胞中迅速降解的蛋白质。因此这些突变的纯合子是 β^0 地中海贫血。

一系列有意义的畸变发生在剪接信号上,同时这些突变大为加强我们对正常的剪接过程的理解,这些剪接畸变中的一些较易于理解,因为它们改变了内含子起始处不变的 GT 序列或末端的 AG,因为这些序列对正常剪接是绝对必要的,毫不奇怪,这些突变导致 β^0 地中海贫血。然而,更精细的突变并不发生在不变的核苷酸而是作用于供体或受体的一致区域。例如图 6-20 所示第一个内含子的供体序列的突变。那些改变了第一个核苷酸,使正常的 GT 序列消失的突变,导致 β^0 地中海贫血,但在内含子中第 5 位和第 6 位的突变使 RNA 正确剪接的能力下降,但仍能导致可检测正常 β 珠蛋白的含量。因此,纯合

状态下的这些突变导致 β^+ 地中海贫血。

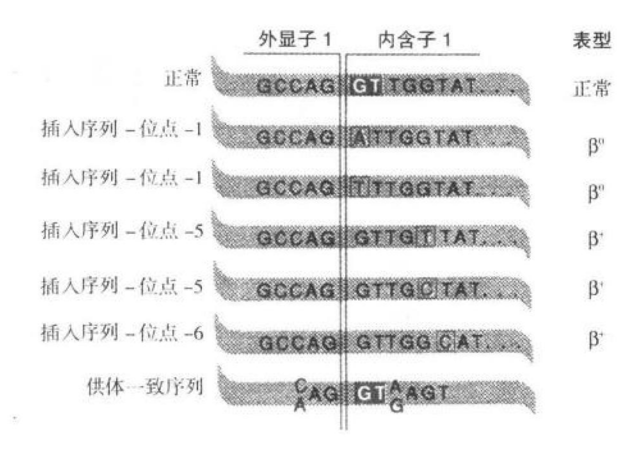
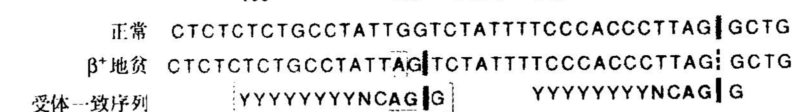


图 6-20 在 β 珠蛋白基因第一个内含子开头部位的五个突变。

一个特别有意义的突变是在地中海地区引起地中海贫血最常见的原因，它是内含子 1 的一个突变，这个突变实际上并未影响正常供体或受体的位点。如图 6-21 所示，这个突变是内含子 1 第 110 位点的 G→A 的单一点突变，发生在正常剪接受体位点上游 21 核苷酸处。此突变的结果是产生了以前在此位点不存在的一个 AG 序列。此 AG 序列之前是一长串嘧啶，而且尽管这个序列并不是理想的适合的一致受体序列，但它的功能极好。如图 6-21 所示，当存在这种突变时，90% 的剪接事件是以新的 AG 作为受体位点，而仅有



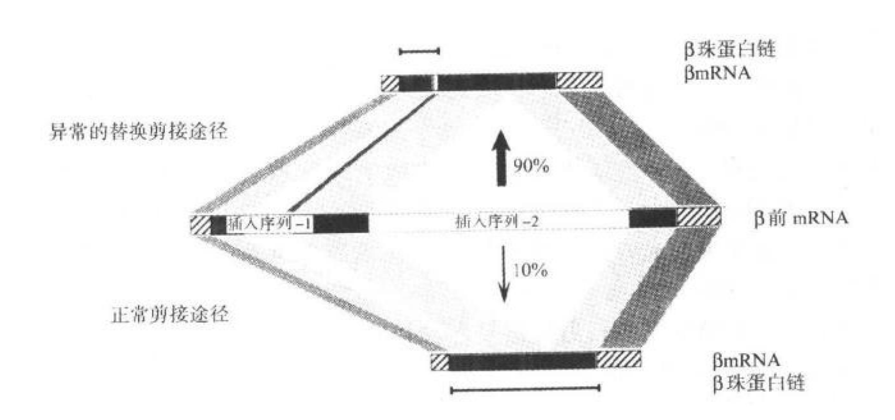


图 6-21 内含子 1 中第 110 位上 G 到 A 的突变导致 β^+ 地中海贫血。这能导致形成一异常剪接受体位点，因为新形成的序列包含 AG 序列，图下部示 90% 用异常剪接途径仅有 10% 的信息正确剪接并产生正常 β 的珠蛋白链，Y = C 或 T；N 为任意核苷酸，一致受体中的不变 AG 有下划线，垂直线表示剪接位置。

10%剪接在正确的位点。此异常剪接的 RNA 含有一个额外的 19 个核苷酸的成熟 RNA，它在正常的剪接过程中已被剪掉。因为这不是 3 的整倍数，其结果是在所有随后的编码区中产生移码突变。所以异常剪接的 RNA 不产生有用的蛋白质，而且仅有正常剪接的 10%的 RNA 是有用的。因此在纯合子时这种表型为 β^+ 地中海贫血。

一种更奇怪的突变是血红蛋白 E，它是东南亚的某些地区中约近 30%的人群携带的一种非常常见的变异体。这种血红蛋白病是一种错综复杂的质量和数量的结合异常。正如图 6-22 所示，血红蛋白 E 是由于密码子 26 的一个点突变引起的，它由正常的 GAG 密码子突变为 AAG，导致谷氨酸变为赖氨酸。可以预测这会导致血红蛋白电泳的异常泳动，实际上正是如此。然而，具有这种突变的个体，由这个突变的等位基因仅产生大约 60%正常量的 β 珠蛋白，这种减少不能从蛋白质稳定性的基础来解释。对此突变基因的剪接模式研究后，蛋白质产物这种额外数量下降的原因也就清楚了。正如图 6-22 所示，密码子 24～27 的序列实际上是非常适合作为剪接的一致供体信号。当然，在正常情况下全然不用这个潜在的供体信号。在理想的内含子供体上，GT 紧随其后的核苷酸既能是 A 或 G，但 A 在此位置的出现与 G 相比通常大约是其 2 倍。显然 β^E 突变是在此位置由 G 变为 A，激活了这个"隐蔽"的剪接位点，以致它开始以一种颇为有效的方式发挥功能。这一事件的结果是约 40%的这种剪接移掉从密码子 25 至外显子 1 末端的序列，或者说移

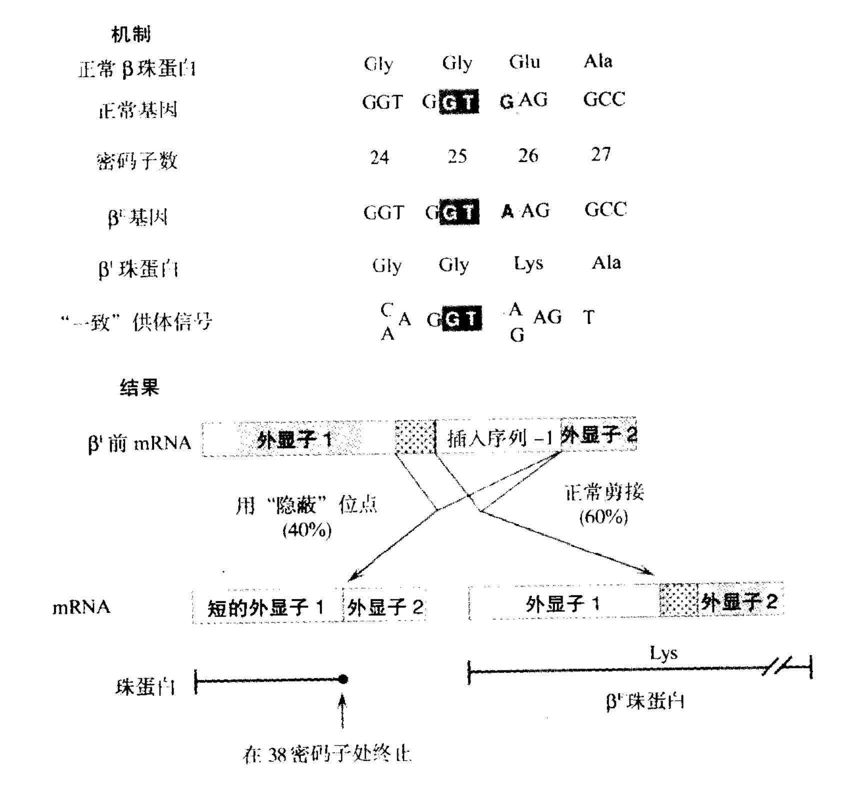


图 6-22 血红蛋白 E 的突变。这种突变产生由于激活隐蔽的剪接供体位而引起数量异常和由于正常剪接的信息在 26 密码子产生赖氨酸而不是谷氨酸引起了质量的异常。密码子 26 上的 A 显然可激活隐蔽的剪接供体位点，40%的剪接事件用此位点。

掉总共16个核苷酸。再者，因为这不是3的整倍数，这一异常的剪接信息是在第2外显子处框码之外，导致在框外的终止密码子处停止。因此，这40%的异常剪接的RNA不产生任何可检测的β珠蛋白。正常剪接的RNA包括β^E突变。因此，这个结果是由单一突变产生的质量和数量结合的异常。

最后一种有意义的突变如图6-18所示是发生在多聚腺苷酸化信号的点突变，即其中AATAAA变为AACAAA。尽管转录能正常发生，但对于正常加工及转运到细胞质中必要的切割及添加A残基却受到损伤，因为最终结果正是β^+地中海贫血表型。

六、β地中海贫血的临床表现

自从Dr Thomas Cooley第一个描述这种疾病后，严重型地中海贫血也称为Cooley贫血，在出生时症状常不明显，因为从胎儿到成人血红蛋白的转换仍未完成，同时β珠蛋白链的缺乏尚未引起后果。然而，在出生后的第一年中，胎儿血红蛋白产量持续下降，严重的贫血症状明显。由于这种贫血，大部分的红细胞在被释放到循环前在骨髓中被破坏，这对血液生成产生重要影响。骨皮质变薄会导致病理性断裂，以及到颜面骨和颅骨的变形。肝和脾也显著增大并成为红细胞再生的另一场所(图6-23)。如不治疗，通常在十岁以前由于严重的贫血，虚弱和感染而发生死亡。




图6-23　严重型地中海贫血的临床图片。左侧的个体18个月大，未接受治疗，她有显著增大的肝脏和脾脏(在她的腹部标出它们的边界)，可以刺激血细胞生成。右侧的个体用输血治疗，外表正常，她戴着输入泵，可以持续泵入去铁胺以解决铁的超负荷，要不然铁超负荷对于输血治疗来说是一个致死性的并发病(引自 Weatherall DJ, Clegg JB. The thalassemia syndromes. 3rd ed. Oxford: Blackwell Scientific Publications, 1981.)。

这些症状通过输血能有效地改善,因为它可供应正常的红细胞且抑制过度活跃的骨髓。然而,归根到底,输血还是一把双刃剑。其原因与铁的代谢机制相关。在正常情况下,人体中铁的总量是在吸收水平来完全调节的。当铁在体内贮积时无有效的机制来提高排泄。在输血过程中,体内铁的总水平持续上升,而且无法排泄,铁沉积于心脏、肝脏、胰脏及其他器官,这会导致这些器官的功能衰竭,特别是心脏,以及一个不可恢复的功能下降过程。图6-24表示用输血治疗严重型地中海贫血患者的存活曲线。虽然在头十年中存活较为正常,此后铁超负荷的效应变得明显,导致在十几岁或二十多岁时死亡。

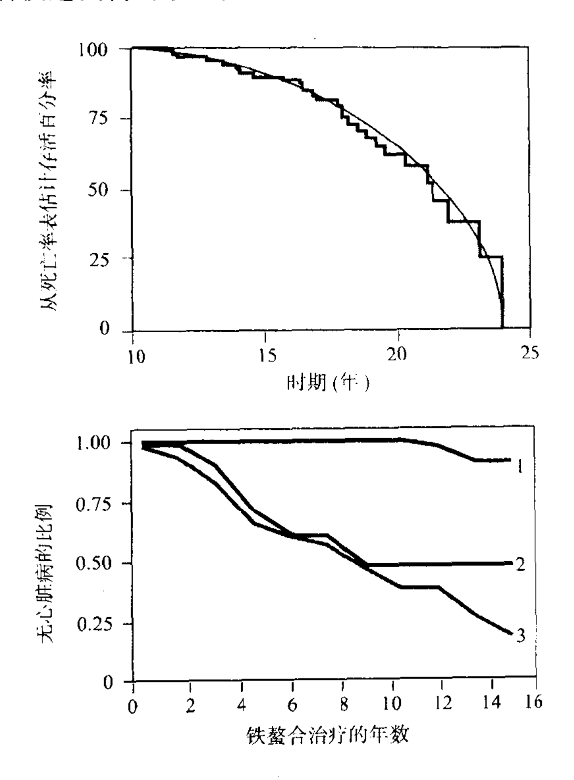


图6-24 地中海贫血的输血及铁螯合治疗。上图示有β地中海贫血症状的个体用输血治疗但未用铁的螯合剂治疗的存活曲线(引自 Weatherall DJ, Clegg JB. The thalassemia syndromes. 3rd ed. Oxford: Blackwell Scientific Publications, 1981.)。当有刺激铁排泄的物质的介入,曲线上生存得以改善,用去铁胺治疗减少过量的铁使由于铁过量而致的并发症减少,包括下图所示的心脏病。1线示病人中无心脏病的存活者用最有效的螯合治疗(<33%的血清铁蛋白量>2500ng/L,铁蛋白是总体铁贮存量的实验室指标)。2线示具有铁蛋白量持续增高的患者中的存活者(33%~67%的铁蛋白量>2055ng/L),3线示大部分铁蛋白量显著增高的病人中的存活者(67%>2500ng/L)(引自 Olivieri NF, Nathan DG, MacMillan JH, Wayne AS, Liu PP, McGee A, Martin M, Koren G, Cohen AR. Survival in medically treated patients with homozygous β-thalassemia. N Engl J Med 1994;331:574~578.)。

利用与铁螯合且使铁从尿中排泄的药物给此情况下的患者带来一些希望。最常用的药物去铁胺看来能够螯合足够的铁,以维持慢性输血,并只要通过静脉泵持续皮下给

药。螯合治疗明显减少包括心脏病在内的铁过量而来的并发症(图 6-24)。然而,常规治疗方法难以忍受而且痛苦,迫切需要能够口服的有效铁螯合剂。

对于镰状细胞贫血,骨髓移植具有潜在的治疗能力,尽管此方法也可有潜在的严重致病率及死亡率,然而已发现这种方法更广泛地应用于地中海贫血,用于一贯预后不良的患严重疾病未经治疗及常规螯合治疗受限的患者。最近研究表明,用这种治疗方法可治愈 80%以上的严重 β 地中海贫血患者。在选择的高危人群中,大规模人群筛查及产前诊断对于严重型 β 地中海贫血发病率也有重要的影响。图 6-25 表示在撒丁岛上 β 地中海贫血纯合子的出生率。现在每年仅生 4~5 例,而与 1975 年相比每年大约 100 例。

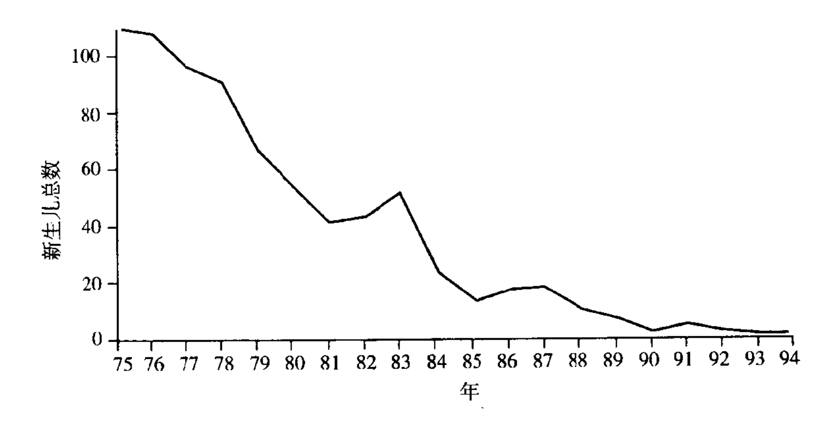


图 6-25 在 Sardima β 地中海贫血纯合子的出生率下降(引自 Cao A, et al. Clinical experience of management of thalassemia: the Sardinian experience. Sem Hematol 1996;33:66~75.)。

七、胎儿血红蛋白的遗传持续增多症

在本章中讨论的最后一组突变是血红蛋白产物的发育时限的变化。遗传性胎儿血红蛋白持续增多症(HPFH)的患者在成年时仍持续大量产生血红蛋白 F(超过总量的 1%),却没有其他原因。在一些病例中这些患者甚至可能产生 100%胎儿血红蛋白。有些惊奇的是,这是一种良性病变,这些的个体并未意识到血红蛋白的异常。人们可推测胎儿血红蛋白,由于它多少能增加氧亲和性,如果他的产物在成人一生中持续增加则产生症状,但这些非常轻微,而不致引起明显的临床疾病。

HPFH 有双重意义。第一,在发育时限中基因突变可提供基因表达一般性调控的线索,这有助于我们加强对本领域的理解。第二,对于胎儿珠蛋白基因表达调控理解的能力可使我们对这种情况下患镰状细胞贫血或 β 地中海贫血的个体进行操作,使胎儿珠蛋白基因返回给这些患者将可能起治疗作用。

HPFH 的分子基础被证实是异源性的(图 6-26)。许多这样的个体具有 δ 和 β 珠蛋白的大片段缺失,其缺失向 3′端方向延伸相当长的距离。然而,胎儿 β 珠蛋白基因并没有缺失,而且以一种不太清楚的机制在整个生命过程中持续以高水平转录,而不在出生之

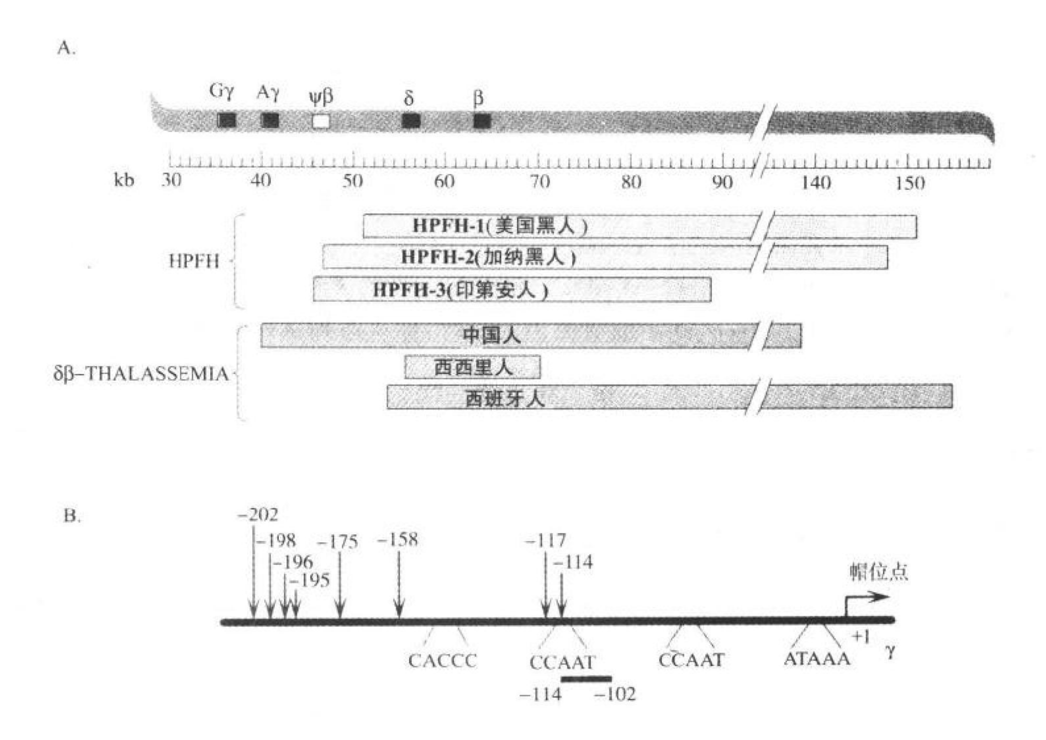


图 6-26 引起 HPFH 的分子机制。在(A)示三个 HPFH 缺失导致包括成人基因在内的 β 珠蛋白基因簇 3′末端 DNA 片段的大的缺失。并不是所有的缺失都导致 HPFH，图中示 3 个 δβ 地中海贫血突变。在(B)示在 Gγ 或 Hγ 胎儿珠蛋白基因启动子中突变的位点(相对于加帽位点而排序)。这些突变在 HPFH(在线上用箭头表示)患者中被发现。这些位点可能位于重要的调控序列，但大部分不在通常的启动子元件(CACCC，CCAAT 或 TATA)。序列下的条框示在一个患者中发现的小的缺失，它去掉了第二个 CCAAT 盒。

时关闭掉。然而不是所有移去 δ 和 β 珠蛋白基因的大缺失都具有这种表型。图 6-26A 也表明胎儿基因中按正常方式关闭的相同类型的三个突变。因为在受累的染色体上 δ 和 β 基因正在消失，因此表型为 δβ 地中海贫血。为什么一些缺失导致 HPFH 而另一些导致地中海贫血的原因仍不知晓。

另一些 HPFH 个体在 β 基因簇中没有主要的缺失，一般来说，这些个体过度表达 Gγ 或 Aγ，但并不是两种都过度表达，而且从他们受累的染色体仍能产生大量的 β 珠蛋白。克隆和测序研究已证明这些个体在过表达的胎儿珠蛋白基因启动子区携有点突变，同时，这些突变至今已被鉴定，如图 6-26B 所示。正如所见，这种突变散布于从 5′到转录起始点的区域，并且预测鉴定的测序涉及正常转换的机制。这些突变之一位于转录起始点 5′114bp 处，发生在重复的 CCAAT 盒之一的第二个 C 上，同时在另一个患者的小缺失去掉了这个序列的大部分。大部分其他突变位于以前肯定是对调控很重要的启动子区域。一种假说认为这些突变的机制是突变位于阻遏物的结合位点，此阻遏物在正常时封闭胎儿珠蛋白的产生，但在存在突变时阻遏物不能与之结合。这个假说仍在积极研究中。从 HPFH 突变的分析中所学到的知识对于试图开发治疗镰状细胞贫血和 β 地中海贫血的有效途径具有很好的临床价值。

八、总 结

在本章我们学习了人类珠蛋白基因中发生的大量突变事件，目前这个系统是人类一系列基因座中研究最为透彻的，由于我们对分子遗传学认识的提高，这一系列丰富的突变事件可预测许多其他基因的发现。对于遗传学的学生来讲，这些突变的详细内容并不特别重要，然而重要的是其体现的原理及了解这些突变的发生机制对正常基因功能所具有的深刻含义。正如 Bateson 70 年前所提倡的"珍惜你所遇到例外的情况"。

推 荐 读 物

普通

Bunn HF, Forget BG. Hemoglobin: molecular, genetic, and clinical aspects, Philadephia: WB Saunders 1986. *A superb and thoroughly referenced work on human hemoglobin, from protein structure and function through molecular biology.*

Dickerson RE, Geis I. Hemoglobin: structure, function, evolution, and pathology. Menlo Park, CA: Benjamin / Cummings, 1983. *A beautifully illustrated short volume on the structure of the hemoglobin protein.*

Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus H. The molecular basis of blood disease. 2nd ed. Philadelphia: WB Saunders, 1994. *Aslo a superb reference text. Chapters on hemoglobin molecular biology and hemoglobin switching are particularly relevant.*

Weatherall DJ, Clegg JB. The thalassemia syndromes. 3rd ed. Oxford: Blackwell Scientific Publications, 1981. *Encyclopedic work on the clinical and laboratory aspects of thalassemia. Most of the information in the book predates the expolsion of information made possible by recombinant DNA, however.*

Weatherall DJ, Clegg JB, Higgs DR, Wood WG. Hemoglobinopathies. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, Inc., 1995. *An up-to-date review by one of the founders of the field.*

经典参考文献

Ingram VM. A specific chemical difference between the globins of normal human and sickle cell anemia hemoglobin. Nature 1956; 178: 792

Pauling L, Itano H, Singer SJ, er al. Sickle cell anemia : a molecular disease. Science 1949; 110: 543

第七章　人类疾病的生化和分子遗传学

以往的经验表明，大自然经常通过其错误赋予我们机会意外地洞察她的许多奥秘——否则它们将是一个封闭的领域。

——A. Loewy 和 Neuberg（引自 Uber Cystinurie. Hoppe-Seyler's Physiol Chem 43:338 ~ 354, 1904）

无论你在研究什么，如果你发现一个突变体你的前景将会更乐观。

——匿名

前一章详细地分析了人类血红蛋白分子的分子病理学，证实了引起人类遗传性疾病的突变机制是富于多样性的，并且介绍了人类疾病的分子遗传学原理。在这一章中，我们将讨论在生化和分子水平已充分了解的另外几种遗传性疾病。这里所举的例子并不详尽，而是选择一些特殊的疾病来证明医学遗传学的特殊原理。Archibald Garrod 对尿黑酸尿症的开拓性研究开辟了生化遗传学这一领域，并且提出了"先天性代谢缺陷"（inborn errors of metabolism）这一概念。他对生化遗传学原理的洞察力是在仔细研究病人的基础上得来的，这是临床研究的一个范例。药物遗传学的建立则是人们偶然观察到药物诱发性贫血和葡萄糖-6-磷酸脱氢酶缺乏症特征的结果。α_1 抗胰蛋白酶缺乏症说明，一种非酶蛋白（蛋白酶抑制剂）的单个突变与吸烟这种最常见的环境因素相互作用，怎样使人易患一种常见的严重疾病——肺气肿。对家族性高胆固醇血症卓越的研究美妙地勾勒出基础遗传学与临床医学之间的动态相互作用，证实了受体介导的胞吞作用一方面是正常的细胞活动，另一方面也是冠状动脉疾病的一个主要原因。对血友病的研究为深入了解人类突变的基本分子机制提供了重要的研究方向。最后，对胶原疾病的分子分析正在加深我们对"蛋白自杀"（protein suicide）和结构蛋白显性遗传性缺陷引发疾病的机制的理解。

一、先天性代谢缺陷

1908 年，Archibald Garrod 在皇家伦敦医学院发表了题为"先天性代谢缺陷"的著名报告，公布了他对四种人类罕见疾病——尿黑酸尿症、戊糖尿症、胱氨酸尿症和白化病的研究结果。尿黑酸尿症以黑尿和脊柱及大关节的退行性关节炎为特征。尿刚排出时实际上是无色的，但与空气接触后，其中大量的尿黑酸被氧化，尿液迅速变成黑色。内源性尿黑酸自身氧化形成的产物沉淀于软骨和胶原组织，使上腭出现蓝色或黑色的色素斑块，

眼巩膜出现黑色沉淀，并且大关节和脊柱椎间盘发生退化性改变(图 7-1)。这些软骨沉淀物在光学显微镜下呈赭石色，所以这种类型的关节炎被称为褐黄病。

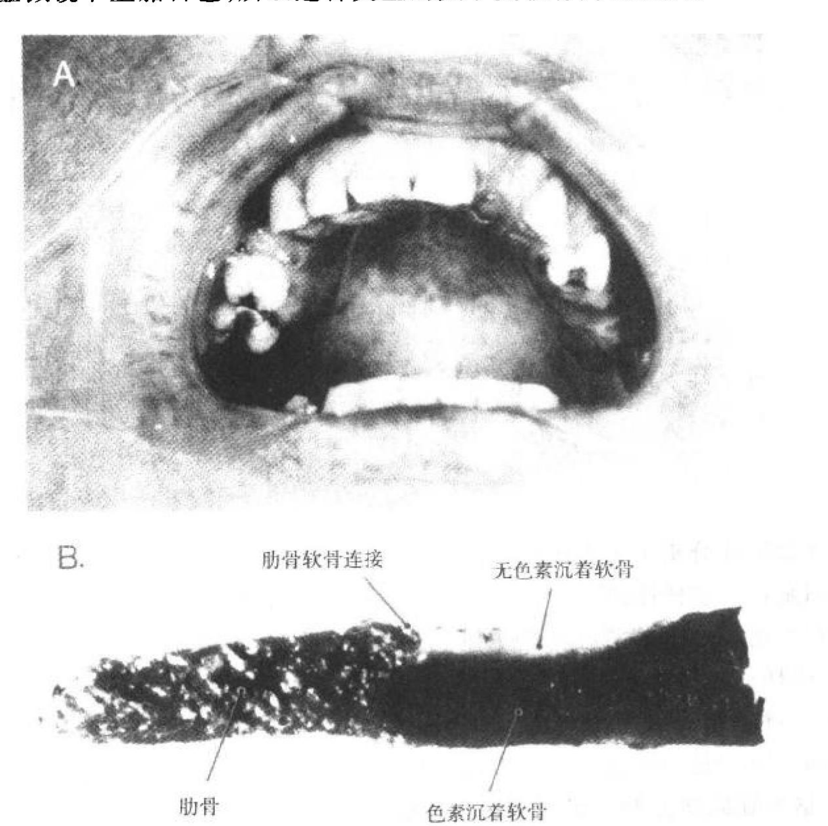


图 7-1 组织中的尿黑酸氧化产物沉着。A 一位尿黑酸尿症患者硬腭的色素沉着。B 尿黑酸尿症患者深度色素沉着的肋软骨(引自 McKusick VA. Heritable disorders of connective tissues. 4th ed. St Louis: Mosby, 1972.)。

Garrod 发现将尿黑酸让受试者服食，尿黑酸尿症患者由尿中定量排出，而正常受试者则不排出。而且，当给予尿黑酸尿症患者高蛋白饮食特别是给予苯丙氨酸和色氨酸时，内源性尿黑酸的排出增多，而给予正常个体这些氨基酸则不会排出任何可检出的尿黑酸。Garrod 分析了这些现象，认为尿黑酸是苯丙氨酸和色氨酸代谢的正常代谢产物，它迅速转化为另一种代谢产物，不在体内蓄积。但尿黑酸尿症患者的下一步代谢被阻断，尿黑酸在体内贮积并大量由尿中排出。他的假设在 50 年后被证实，即尿黑酸尿症患者的肝脏中检测不到尿黑酸氧化酶，此酶负责尿黑酸的进一步代谢(图 7-2)。Garrod 通过他对尿黑酸的研究概括道：代谢的每一步都有独特的酶起作用。他在 1908 年写道："整体代谢的概念正在让位于分步代谢。这个观点每天都获得证据，即蛋白和糖的每一成分合成和分解的多个连续步骤都有特殊的酶为其特定的目的起作用。"

尿黑酸尿症的家族分布也给 Garrod 以深刻印象。他注意到这种异常"……易于发生在那些父母并不表现异常的几个兄弟和姐妹中，而由父母直接传给孩子的很少。"

在 17 个尿黑酸尿症的家庭中有 8 个父母是一级表兄妹。Garrod 请教了他的朋友、著名的英国生物学家 William Bateson(他创造了“遗传学”这个术语)。Bateson 根据新近重新发现的孟德尔定律对这种家族模式作出解释,指出一级表兄妹的婚配恰好最有可能为一种罕见的隐性性状表现出来提供条件。因此,尿黑酸尿症是第一种被确认的常染色体隐性遗传病。这种认识意指正常的等位基因是未受累个体特异的酶产物所必需的。这是关于基因可能通过编码的酶施加其影响的最初的线索。Garrod 的工作预言了一个基因一种酶的概念。这个概念以 George Beadle 和 Edward Tatum 为先驱,创立于 20 世纪 40 年代和 50 年代。Beadle 和 Tatum 对链孢霉菌(mold *Neurospora*)进行了研究,用紫外光和 X－射线来获得营养型和营养缺陷型的突变株。这种突变株缺乏催化特定代谢步骤的能力。他们证实链孢霉菌中每个代谢步骤都由某些酶控制,单个突变改变单个功能。Beadle 和 Tatum 完成他们的工作时并不知道 Garrod 的发现,但 Beadle 在他 1958 年接受诺贝尔奖的讲演中称赞了 Garrod:“……我们再次发现了 Garrod 很久以前就认识到的东西,现在我们了解了他的工作,并且知道在原理上我们并没有增加什么新的东西。”

图 7-2 苯丙氨酸和酪氨酸代谢为乙酰乙酸过程中的酶催化步骤。

最后,Garrod 确认了生化和遗传个性的概念:“化学个性的存在是化学特性的必然后果。甚至那些特异体质者对药物和食物成分的过敏反应(如谚语说有些东西对一个人来说是肉,对另一个人来说可能是毒药),可能也有其化学(和遗传)基础。”这样,Garrod 为下面将要讨论的药物遗传学的研究奠定了基础,还对第四章讨论过的部分由外源性因素引

起的疾病的遗传基础建立了概念性的框架。

Garrod的观点大大超前于他的时代。尽管他在英国医学界有着很高的地位(他继承了 William Osler 在牛津的皇家医学院教授职位),他在先天性代谢缺陷方面的工作对临床医学和遗传学并未立即产生多大的影响。他预言将会发现许多先天缺陷,提出存在"几乎无数种这样的先天畸形"。实际上,以后的许多年并没有发现更多的代谢先天缺陷,很可能因为严重的先天缺陷在婴儿早期就导致死亡,而且在这段时间感染性疾病是婴儿死亡的主要原因,所以先天缺陷还没有被认识。

对先天性代谢缺陷认识的突破不得不等到抗生素应用使得婴儿感染性疾病的死亡率显著降低,以及营养和诸如婴儿腹泻等疾病得到改善以后。此外,某些技术的进步也非常重要。纸层析法的建立可以将尿液和血液中氨基酸与其他代谢产物分开;人类细胞培养技术的建立使得研究人员可以研究病人的组织而不需经常地与病人接触;随后分子技术的发展则使人们可以利用更精尖复杂的方法研究人类的生化疾病。

二、药物遗传学

在朝鲜战争期间,美国士兵使用伯氨喹啉来预防疟疾。约10%的黑人士兵发生血管内红细胞破坏(溶血)并继发急性的但是自身限制的贫血。少数白人士兵(通常是地中海血统)患上一种相似的但严重得多的溶血性贫血。后来发现这种药物诱发性贫血的基础是葡萄糖-6-磷酸脱氢酶(G6PD)的遗传性缺陷。此酶催化磷酸己糖旁路的第一步。葡萄糖代谢的这个途径的一个重要功能是产生 NADPH,后者对于保持谷胱甘肽和巯基水平以保护细胞蛋白免受氧化剂损害是必需的(图 7-3)。

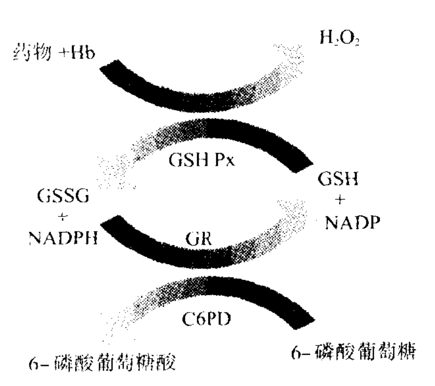


图 7-3 G6PD 在 NADPH 生成中的作用(NADPH 用于防护血液红细胞免于氧化损伤)。NADPH 在谷胱甘肽还原酶(GR)催化的氧化型谷胱甘肽(GSSG)还原为还原型谷胱甘肽(GSH)的过程中是必需的,后者在有过氧化物存在时被谷胱甘肽过氧化物酶(GSH Px)氧化。

大约10%的黑人 G6PD 电泳迁移率有改变,并且只有正常酶活性的15%。这些人临床表现完全正常,除非使用某些药物如抗疟疾药伯氨喹啉、磺胺或呋喃妥英类抗生素,或是患某些感染性疾病如病毒性肝炎或细菌性肺炎。在这些环境因素下,他们发生急性溶血性贫血。受累个体的新鲜红细胞较陈旧红细胞有更强的抗氧化剂损害能力和更高的 G6PD 活性。随后有人发现,异常的 G6PD(标示为 A-)在体内稳定性降低,半衰期从62天降至13天。尽管其他细胞类型也有相同的缺陷,但临床表现局限于贫血。这是因为成熟的红细胞缺乏细胞核和核糖体,所以不能合成新的 mRNA和蛋白来替代被降解的蛋白质。这样,氧化压力就会选择性地破坏 G6PD 活性减少的红细胞。然而,健康个体的骨髓能够对这种压力作出反应,增加红细胞的生成量。这些新鲜红细胞由于具有正常的 G6PD 水平,可以抵抗氧化压力。所以,只要骨髓能够增加红细胞产量,即使诱发药物持续存在,这种溶血性贫血也是自身限制的。

在一些地中海 G6PD 缺乏患者尤其是小孩中,食用蚕豆(地中海地区一种常见的扁豆)会引起非常严重的急性溶血危象(蚕豆病),常常需要输血。早在公元前五世纪,希腊历史学家 Herodotus 就曾描述了一些希腊人食用蚕豆后的异常反应。所以,G6PD 缺乏症是 Garrod 预言的一个绝好的证据,即特异体质者对药物和食物成分的反应是先天性代谢缺陷的结果。

家系研究表明 G6PD 由 X 染色体上的一个基因编码。随后,这个基因被定位于 X 染色体长臂末端,在甲型血友病基因座(下文讨论)与色彩视觉基因座之间。G6PD 基因全长约 19kb,含 13 个外显子,编码一个含 514 个氨基酸的蛋白质。

G6PD 突变型最初被发现是因为在淀粉凝胶中的电泳迁移率不同。G6PD B 是大多数人群中的野生型,G6PD A 是黑人人群中常见的突变型(见于 10%男性非裔美国人),如此命名是因为后者电泳迁移率快得多。G6PD A 具有全酶活性,而上面讨论过的 G6PD A - 电泳迁移率与 A 相近,但只有 G6PD A 或 B 酶活性的 15%。

最近的研究探明了上述这些以及其他 G6PD 突变型的分子缺陷(图 7-4)。第五外显子的 376 位核苷酸发生 A→G 改变,导致 126 位氨基酸由天冬酰胺替换为天冬氨酸,这种替换作用可能就是 A 型及 A - 突变型电泳迁移率比 B 型或野生型 G6PD 快得多的原因。G6PD A - 发生二次突变,即外显子 4202 位核苷酸的 G→A,导致 68 位氨基酸由缬氨酸替换为蛋氨酸,这可能是这种突变型稳定性降低的原因。

在 G6PD 的地中海突变型中,酶活力几乎检测不到,发生溶血性贫血的危险较大。其外显子 6563 位核苷酸 C→T,导致 188 位氨基酸由丝氨酸替换为苯丙氨酸,这可能是这种突变型酶催化活力和稳定性降低的原因。

在 DNA 水平已报道了 60 多种独立的突变,范围包括几乎整个编码区(图 7-4)。令人吃惊的是,几乎所有的突变都是错义突变,而且大多数是编码链或非编码链中 CpG 双核苷酸中发生 C→T 的转换突变。CpG 双核苷酸通常被认为是突变的"热点",因为胞嘧啶经常发生甲基化,而 5 - 甲基胞嘧啶可自发地脱氨基转变为胸腺嘧啶。这些突变的表型是高度可变的,大多数无症状,但有些即使在没有诱发性感染或药物存在时也会引起慢性溶血性贫血,还有一些会导致严重的新生儿黄疸和核黄疸(胆红素沉积继发的脑损害)。这些严重得多的突变大多集中在外显子 10 靠近 NADP 结合域的地方。

已在蛋白质水平报道了四百多种等位突变型,其中许多发生频率是多态的。大多数突变型等位基因是根据其被发现时的地域命名的,依字母顺序从西德的亚琛(Aachen)排列到乌克兰的 Zhitomir。

G6PD 缺乏症很常见,全球大约有四亿人受累(图 7-5),在非洲、地中海及亚洲人群中有较高的发病率。由于这些地区疟疾流行,所以有一种观点认为,对镰形疟原虫疟疾的抗性可能是这种平衡多态的原因。G6PD 缺乏症的发生率在一些犹太人群中超过 60%,在撒丁岛的村民中超过 30%,沙特阿拉伯人为 13%,男性非裔美国人中超过 10%。世界上不同地区的 G6PD 缺乏症由多种不同的突变型引起。与其说这提示人群迁移传递一种有利的突变,不如说是趋同进化。尽管 G6PD 缺乏症杂合子看来并没有什么优势,但发现 G6PD A - 突变型女性杂合子体内寄生虫数较少,且对恶性疟疾的抵抗力相对较强。

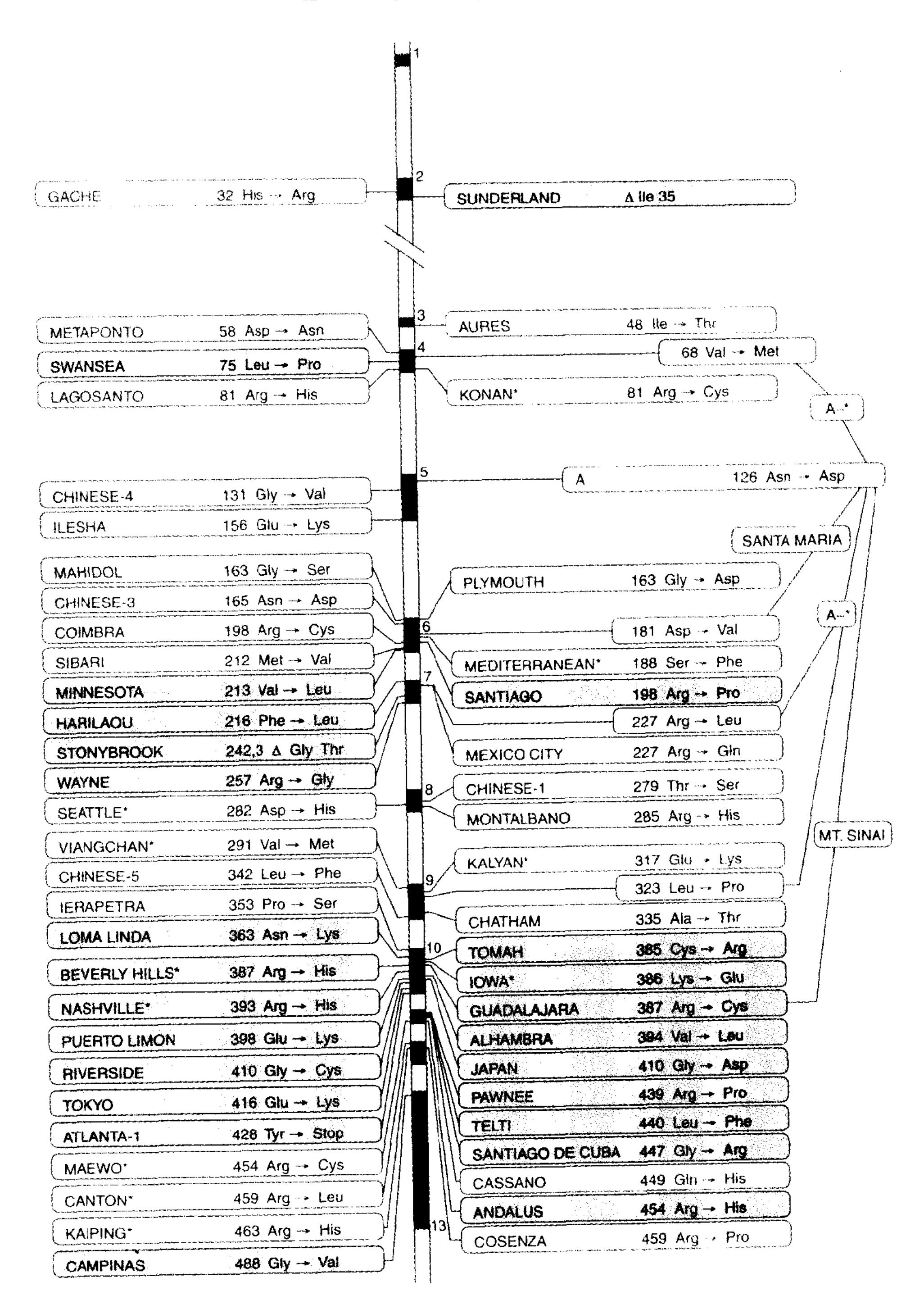


图 7-4　人类 G6PD 基因的突变 不同名称变异型中的氨基酸替换见 G6PD 基因简图。外显子用标有数字的黑框表示。有灰色背景的变异型与慢性溶血性贫血有关。星号表示其他变异中有相同的突变。G6PD Mount Sinai 和三种不同类型的 G6PD A－中有两种突变(引自 Vulliamy T, Beutler E, Luzzatto L. Variants of glucose-6-phosphate dehydrogenase are due to missense mutations spread throughout the coding region of the gene. Hum Mutat, 1993;2:159～167.)。

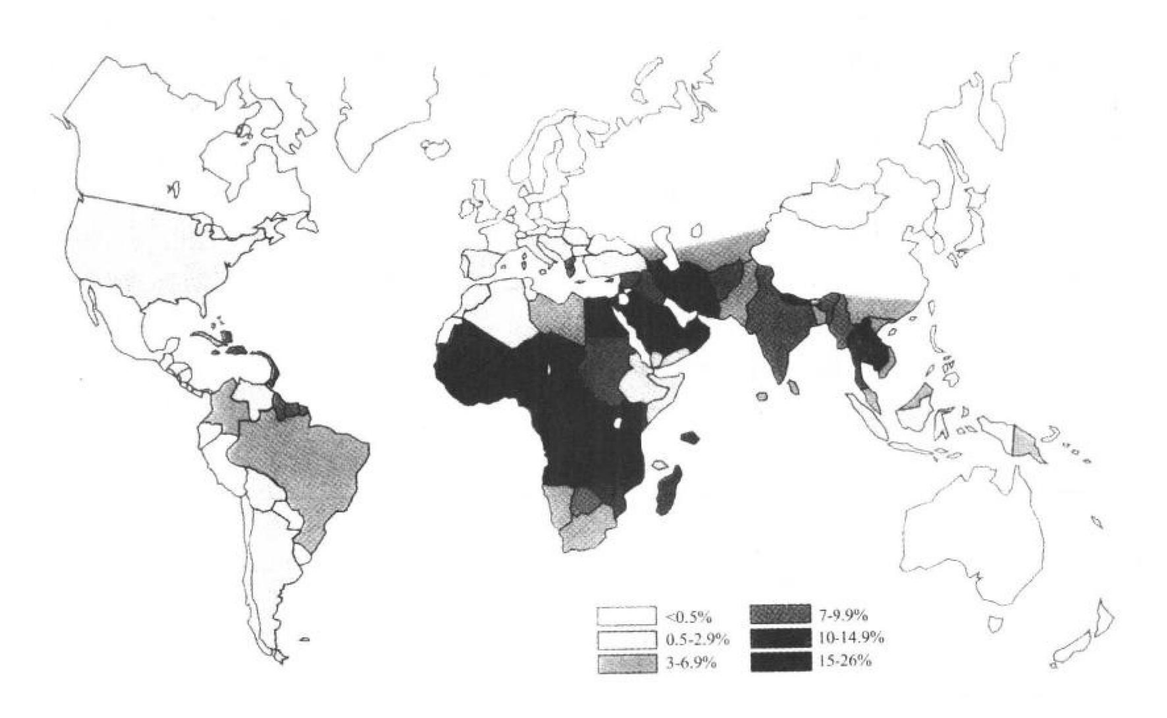


图 7-5 G6PD 缺乏症的世界分布及大致人口频率（引自 Luzzatto L, Mehta A. Glucose-6-phasphate dehydrogenase deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease.7th ed. New York: McGraw-Hill, 1995:3367～3398.）。

G6PD 缺乏症是药物遗传学(研究对药物反应差异的遗传学基础)的一个范例。对另外几种特异影响药物代谢酶的多态性也有论述,它们可能是不同个体之间药物疗效差异以及对药物的特应性反应的原因。药物代谢性酶在不同人种之间的遗传学差异则可能是药物疗效的种族差异(例如,东方人种和高加索人种之间 β 阻断剂疗效的差异)以及东方人对酒精诱导的面红及心悸更为易感的原因。Garrod 的化学个性概念也提示,对外源性病原体的反应可能有遗传差异,它们可能是引起不同疾病易感性的原因。对此,α_1 - 抗胰蛋白酶缺乏症提供了一个很好的例子。

三、α_1 - 抗胰蛋白酶缺乏症

丝氨酸蛋白酶是一组紧密相关的蛋白水解酶,其活性位点存在丝氨酸,在凝血和纤溶过程及激肽和补体的激活过程起着关键的作用。这些酶的激活至少部分由一些特异的抑制剂控制,它们总称为丝氨酸蛋白酶抑制剂。在血浆中发现的含量最高的丝氨酸蛋白酶抑制剂是 α_1 - 抗胰蛋白酶,一种 52kDa 的糖蛋白,占血浆中 α_1 - 球蛋白总量的 90%。与其名称不同的是,α_1 - 抗胰蛋白酶的主要功能是抑制肺部中性粒细胞产生的弹性蛋白酶的活性。α_1 - 抗胰蛋白酶缺乏症主要表型为肺泡破坏引起慢性阻塞性肺病或肺气肿。

α_1 - 抗胰蛋白酶基因呈高度多态性,在欧洲人群中存在 75 种以上不同的等位基因。

不同类型的 α_1 - 抗胰蛋白酶一般可通过等电聚焦凝胶上的电泳迁移率不同来区别。经常把蛋白酶抑制剂称作 Pi,突变型根据其向阳极泳动的距离加一个字母。欧洲人群中最常见的等位基因是 Pi^M(实际上是四种等位基因,区别在于一个氨基酸的不同,但其电泳迁移率相同),等位基因频率为 0.95。90%的欧洲白人基因型为 MM。两种突变型等位基因 S 和 Z,是大多数与 α_1 - 抗胰蛋白酶缺乏有关的疾病的原因。重要的 *Pi* 基因型的频率以及对应的酶活力见表 7-1。Pi^{ZZ}具有 10% ~ 15%的正常酶活力,在北欧血统的白人中大约为 1/2 500。这种突变型是大多数与 α_1 - 抗胰蛋白酶缺乏有关疾病的原因。纯合子 Pi^{SS}只降低 α_1 - 抗胰蛋白酶活性 50% ~ 60%,不引起疾病。但杂合子 Pi^{SZ}的个体只有 30% ~ 35%的正常酶活性,可能患肺气肿。此外,据报道 12 种以上的罕见等位基因产生严重的 α_1 - 抗胰蛋白酶缺乏或无可检测 α_1 - 抗胰蛋白酶(无效等位基因)。

表 7-1　临床上重要的 *Pi* 基因型

	Pi 基因型				
	MM	MZ	SS	SZ	Z
频率	0.90	0.038	0.001	0.0012(1/800)	0.0004(1/2 500)
活性(%对照)	100	60	50 ~ 60	30 ~ 35	10 ~ 15

患有 α_1 - 抗胰蛋白酶缺乏症的个体患肺气肿的风险至少增加 20 倍,80% ~ 90%的患者最终患此病。激活的中性粒细胞产生的弹性蛋白酶如果不被蛋白酶抑制剂灭活,可以引起肺组织的分解。更进一步说,激活的中性粒细胞释放氧自由基和氯化的氧化剂,这些氧化剂可氧化 α_1 - 抗胰蛋白酶活性部位的蛋氨酸。这种氧化作用使抑制剂与中性粒细胞的结合力下降了 2 000 倍,显著地降低了其抑制弹性蛋白酶活性的能力。这种不受抑制的弹性蛋白酶活性被认为是引起肺组织降解的原因。

临床和流行病学的研究表明,α_1 - 抗胰蛋白酶缺乏症引发的严重疾病在吸烟者中比非吸烟者中多得多。其基础可能是吸烟对中性粒细胞和巨噬细胞合成氧自由基的影响。这样,吸烟这种环境因素与 α_1 - 抗胰蛋白酶缺乏这种遗传素质相互作用,引起严重的肺病。α_1 - 抗胰蛋白酶缺乏的非吸烟者也会患肺气肿,但发病一般比吸烟者晚,其预期寿命虽比未受累者相对减少,但并不像吸烟者那样严重(图 7-6)。

Pi^{ZZ}个体还会患肝病,有人认为这是肝细胞粗面内质网中突变蛋白堆积的结果。大约 10% ~ 15%受累个体患新生儿胆汁郁积性肝炎,其中大约 20%发展为青少年肝硬化。约 20%患 α_1 - 抗胰蛋白酶缺乏症的成人也患肝硬化,而且也增加了患原发性肝癌的风险。

α_1 - 抗胰蛋白酶基因在 14 号染色体上大约占 12 kb,编码一个 1.4 kb 的 mRNA 转录物,此 mRNA 编码一个含 394 个氨基酸的蛋白。Pi^S 突变型是由于第三外显子发生 GAA→GTA 突变,导致 264 位氨基酸由谷氨酸替换为缬氨酸,这使抑制剂的稳定性降低。Pi^Z 突变型是由于第五外显子编码 342 位谷氨酸的 GAG 突变为编码赖氨酸的 AAG。这种突变

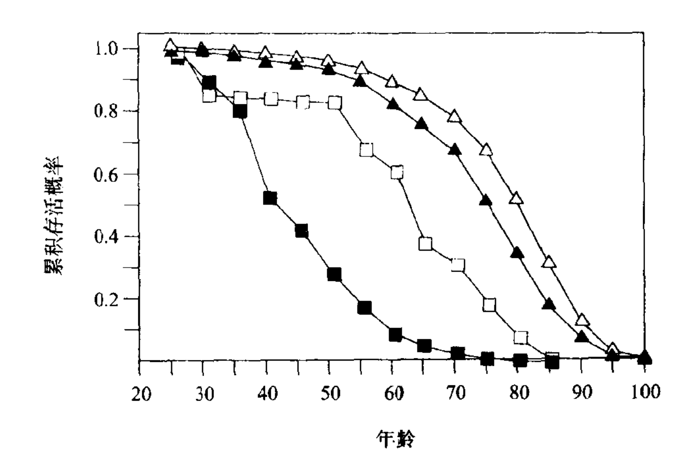


图 7-6 假定已活到 20 岁，吸烟和非吸烟 Pi^{ZZ}个体与全体瑞典人累积存活概率的比较（引自 Larsson C. Natural history and life expectancy in severe α_1-antitrypsin deficiency, PiZ. Acta Med Scand 1987;204:345～351）。

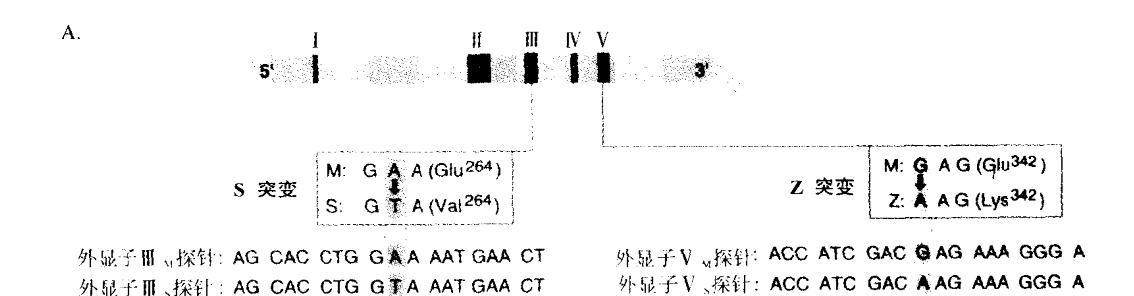
■ 吸烟的 Pi^{ZZ}男性和女性 □ 非吸烟的 Pi^{ZZ}男性和女性
△ 全体瑞典女性 ▲ 全体瑞典男性

改变了蛋白的三级结构，使其在糙面内质网内聚合。其结果是，只有 15%的突变蛋白被肝细胞分泌。此外，这种突变型蛋白作为中性粒细胞弹性蛋白酶的抑制剂的效力比正常型的低。罕见的"无效"突变者在血清内几乎没有可检出的免疫反应性蛋白，这是由于小的插入或缺失导致移码突变从而使翻译提前终止引起的。

尽管二十多种突变引起 α_1-抗胰蛋白酶缺乏，但只有两种突变（Z 和 S）是绝大多数蛋白酶抑制剂缺乏相关疾病的原因。所以，对这种疾病进行产前诊断是可行的。例如，用等位特异性寡核苷酸探针（图 7-7），这种技术在第五章已讨论。曾有人尝试过注射从血浆中纯化或通过重组 DNA 技术获得的人抑制剂。对一系列已确认为呼吸疾病患者的研究表明，每周或每月注射纯化抑制剂，可以将血液和肺泡液中 α_1-抗胰蛋白酶的水平恢复到免受中性粒细胞弹性蛋白酶损伤的水平。很显然，对 α_1-抗胰蛋白酶缺乏症患者，应比对未受累者更强烈地建议他们不要吸烟。将来的替代疗法包括用气雾剂将抑制剂直接运送到肺或选择性地合成弹性蛋白酶抑制剂。肺移植已成功地应用于一些晚期肺气肿病人。

目前对基因疗法也进行了探索（见第十二章）。将人 α_1-抗胰蛋白酶的 cDNA 转染鼠成纤维细胞后，在体外或将细胞植入免疫缺陷裸鼠的腹腔都可合成抑制剂。其他的策略包括用改装的腺病毒携带基因，将其直接导入肺上皮细胞。不幸的是，用纯化的抑制剂替代或用体细胞基因疗法都不能防治 *ZZ* 基因型引起的肝病，除非能关闭内源性突变基因表达。

α_1-抗胰蛋白酶缺乏症在高加索人种中是一种常见的常染色体隐性遗传病（黑人和亚细亚人很少见），许多患者在生命晚期发展为肺气肿或肝硬化。尽管他们在此前没有症状，但如果接触一种特定的环境因素即吸烟，则会有高度的风险。所以，α_1-抗胰蛋白



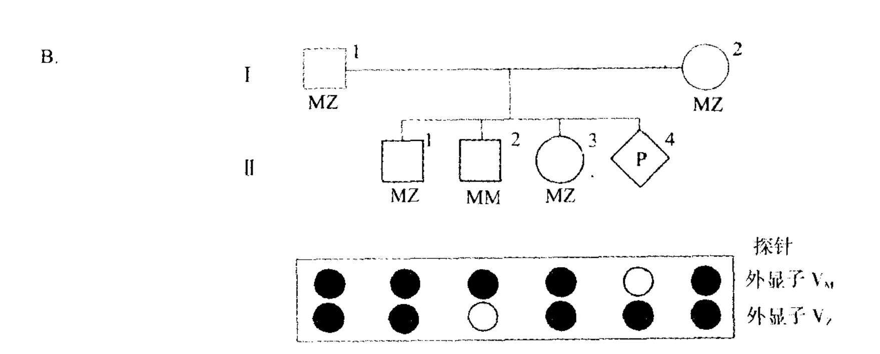


图 7-7　用合成的等位特异性寡核酸探针识别 α_1 - 抗胰蛋白酶基因中 S 和 Z 突变的策略 A. α_1 - 抗胰蛋白酶基因用其 5 个外显子表示（标有 I ～ V 的方块）。基因结构下的方框表示编码 S 和 Z 型 α_1 - 抗胰蛋白酶的突变。图形下方是检测这两种突变的 19 个碱基的寡核苷酸探针（引自 Nukiwa T, Brantly M, Garver R, Paul L, Courtney M, LeCrocq J-P, Crystal RG. Evaluation of "at-risk" α_1-antitrypsin genotype SZ with synthetic oligonucleotide probes. J Clin Invest 1986;77:528 ~ 537.）。B. 用合成的等位特异性寡核苷酸探针对一患 Pi^{ZZ} 型 α_1 - 抗胰蛋白酶缺乏症胎儿进行产前诊断。基因组 DNA 取自系谱中的个体（包括待诊胎儿），用 PCR 进行扩增，扩增的 DNA 加在尼龙滤膜上作为“印迹”，与标记的对 M 和 Z 序列特异的寡核苷酸探针进行杂交，放射自显影检测。用标准电泳技术对待诊胎儿的父母及同胞进行 Pi 分型。Ⅰ-1、Ⅰ-2、Ⅱ-1 及Ⅱ-3 的 DNA 与两种探针都能杂交，证实为 Pi^{MZ} 型杂合子；Ⅱ-2 的 DNA 只与 M 特异探针杂交。胎儿的 DNA 只与 Z 特异探针杂交，表明是突变等位基因的纯合子。

酶缺乏症是“生态遗传”病的一个范例，由一种不利的环境因素和一种特定的遗传素质相互作用而致病。生化和分子的研究揭示了引起 α_1 - 抗胰蛋白酶缺乏症相关疾病的两种常见突变的本质，提供了对此疾病进行高度准确的产前诊断的技术，并促进了替代疗法的发展和体细胞基因疗法的开端。

四、家族性高胆固醇血症

家族性高胆固醇血症（FH）是一种常染色体显性遗传病，在人群中发生率为 1/500，在 60 岁以前发生心梗的患者中大约占 50%。此病的特点是血清胆固醇增高（300 ~ 600mg/dl，正常 < 230）及低密度脂蛋白（LDL）胆固醇增高（ > 200mg/dl）。大约 50%的患者

出现伸肌腱的胆固醇沉积(黄色瘤),比较典型的是 Achilles 腱的不规则增厚以及黄斑瘤,即眼睑的黄色脂肪沉积。更明显的是,患者较早发生心血管动脉粥样硬化。在第三章中(图3-3)所描述的 W.H. 的病史和家族史在此病中很典型。考虑到它是常染色体显性遗传(强调这一点非常重要),诊断一名患者就可以立即鉴别一群高胆固醇血症和早发动脉粥样硬化性心脏病的高危人群。所有一级亲属有50%患病的可能,并应进行检测。高胆固醇血症在儿童中也存在,所以可在早期进行检测,在患者发展为动脉粥样硬化性心脏病以前通过规定食谱和药物加以预防。由于此病相对比较常见,杂合子之间可偶尔发生婚配(1/500 × 1/500 = 1/250 000),其后代为纯合子机率约为百万分之一。纯合子个体血胆固醇非常高(600 ~ 1 200mg/dl),常在儿童期发作心脏病,通常在 20 ~ 30 岁以前死于冠心病。家族性高胆固醇血症的重要不仅在于其常见、严重和可治性,而且因为 Michael Brown 和 Joseph Goldstein 的杰出研究出色地证实了医学和遗传学之间的相互渗透。为此,他们被授予 1985 年的诺贝尔奖。

(一) 家族性高胆固醇血症的细胞缺陷

1974 年,Brown 和 Goldstein 机敏地选择培养的纯合子个体皮肤成纤维细胞作为他们的模型系统,发现这类病人的成纤维细胞存在高水平的 β - 羟 - β - 甲基戊二酸单酰 CoA 还原酶(HMG-CoA 还原酶,胆固醇生物合成的限速酶),在与 LDL 一起孵育时酶活性不像正常那样受到抑制。但是,这种缺陷并不在于对 HMG-CoA 还原酶本身的调节,所以不在于胆固醇的合成,而在于调节信号即胆固醇的传递。他们用放射性碘标记 LDL,证实正常细胞表面对 LDL 的结合是可饱和的。电镜的研究显示 LDL 受体并不是随机分布于细胞表面,而是位于一个含有一种特殊蛋白——网格蛋白、称作有被小窝的特殊部位。LDL 颗粒与 LDL 受体结合后引起随后的内在化,载脂蛋白降解为氨基酸,LDL 颗粒中的胆固醇酯降解为游离胆固醇。细胞内游离胆固醇的氧化衍生物(如 25 - 羟基胆固醇)启动三个重要的调节过程:HMG-CoA 还原酶合成显著减少引起内源性胆固醇合成减少;乙酰 CoA:胆固醇酰基转移酶(ACAT)活性增加引起胆固醇酯化作用和贮存增加;更重要的是,LDL 受体合成减少引起细胞对外源性胆固醇的摄取减少(图7-8)。

家族性高胆固醇血症纯合子患者的成纤维细胞不能结合和内在化 LDL 以及抑制 HMG-CoA 还原酶,所以,家族性高胆固醇血症的缺陷似乎是功能性 LDL 受体的不足。更常见的是患者 LDL 受体部分缺失,如那些家族性高胆固醇血症杂合子,他们有一正常的等位基因编码正常的 LDL 受体,表现为中等缺陷。由杂合子培养的成纤维细胞中,LDL 可对 HMG-CoA 还原酶进行正常调节,但周围 LDL 浓度较正常成纤维细胞所需为高(图7-9)。所以,正常调节依赖于增高的 LDL 胆固醇水平。这些研究除了说明 FH 的分子缺陷,还阐明了一个非常重要的细胞正常活动,即受体介导的细胞内吞。通过这一活动过程细胞特异结合并内吞各种重要的配体,包括激素、生长因子、酶和营养信号。

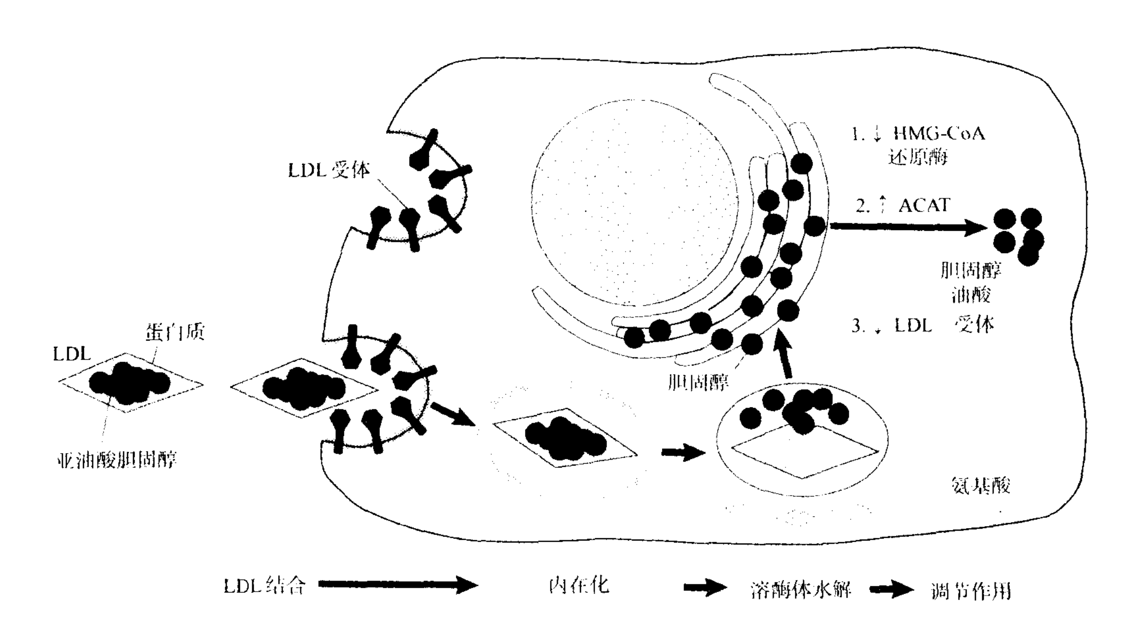


图 7-8 培养的哺乳动物细胞中 LDL 受体途径的连续步骤 HMG-CoA 代表 β-羟-β-甲基戊二酸单酰辅酶 A；ACAT 表示乙酰辅酶 A 胆固醇酰基转移酶；垂直箭头表示调节效应的方向(引自 Brown MS, Goldstein JL. Receptor-mediated endocytosis: insights from the lipoprotein receptor system. Proc Natl Acad Sci USA 1979;76:3330~3337.)。

(二) LDL 受体

纯化 LDL 受体及对其 cDNA 和基因进行克隆，表明受体在内质网合成，在经过高尔基复合体时糖基化，是一分子量约为 160 kDa 插入细胞膜的整合膜蛋白。LDL 受体的结构域经氨基酸序列分析及对其特定突变的研究已经明确(图 7-10)。分子的 N 末端含有配体结合区，该区域有很强的负电荷，由 7 个重复单位构成，每个单位有 40 个氨基酸，其中包括 6 个半胱氨酸。与这个区域相邻的是一个类似表皮生长因子(EGF)前体的区域、一个有 *O*-连接糖的短链，然后是跨膜区和 C 末端一个 50 个氨基酸的胞浆内尾部。后两个区域对锚定受体和将受体适当地定位于细胞膜的有被小窝内非常重要。

LDL 受体基因位于 19 号染色体上，约为 45kb，有 18 个外显子，编码一个 5.3kb 的 mRNA。对这个基因的分析表明，配体结合区由 5 个外显子编码(2~6)，EGF 前体同源区由 8 个外显子编码，糖基化部分由一个外显子编码，跨膜区由第 16 外显子和 17 外显子的一部分编码，而胞浆内的尾部由外显子 17 的剩余部分和外显子 18 的一部分编码。

(三) LDL 受体的突变

在家族性高胆固醇血症中，通过 PCR、单链构象多态和 DNA 测序在分子水平鉴定了 150种不同的突变，并发现有多种突变型和高度的等位基因异质性(图7-11)。大多数家

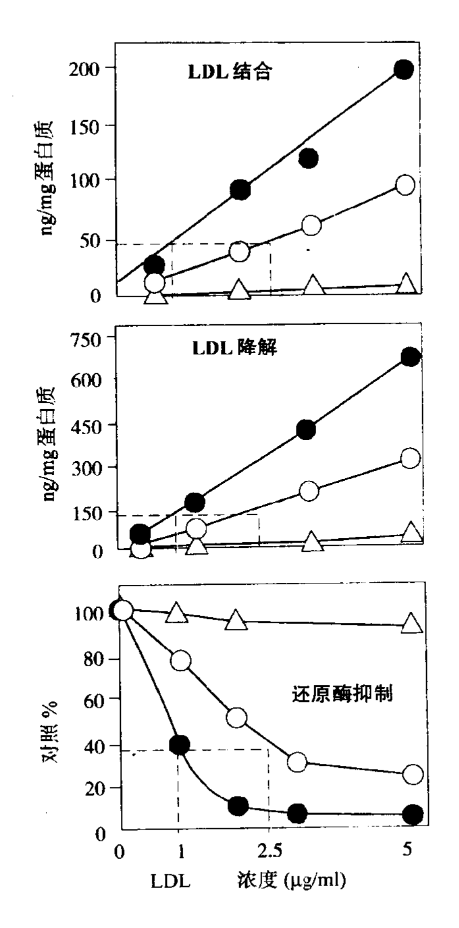


图7-9 家族性高胆固醇血症的细胞缺陷。培养的皮肤成纤维细胞中LDL结合、LDL降解及HMG-CoA还原酶的抑制。

●正常个体 ○FH杂合子

△FH纯合子

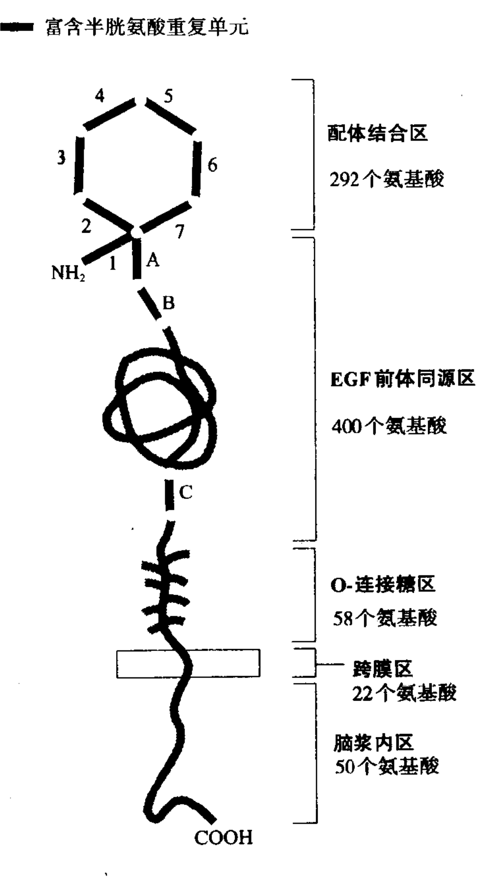


图7-10 人LDL受体的区域结构。富含半胱氨酸的重复单元用标有1~7和A~C的条表示。通常认为重复单元2~7构成主要的LDL结合位点(引自Esser V, Limbird LE, Brown MS, Goldstein JL, Russell DW. Mutational analysis of the ligand-binding domain of the low-density lipoprotein receptor. J Biol Chem 1988;263:13282~13290.)。

系中都有其他家系没有的独特突变,而实际上大多数"纯合子"都被证明是复合的杂合子,即LDL受体基因不同等位基因突变的杂合子。但也有例外,如60%法裔加拿大病人中大的5'端缺失、35%芬兰病人中的9.5kb 3'端缺失、基督(Christian)黎巴嫩病人中第14外显子的无义突变及立陶宛血统的德系犹太人中的外显子4的3bp缺失,因为它们可见于这些人群中的许多家系。有人认为这种情况反映了建立者效应(见第四章)。

分析LDL受体基因大的插入和缺失,可得到一些关于这些突变的可能机制以及LDL受体基因进化的有趣的资料。对一个外显子2有14kb插入的家系的分子分析表明,这种突变引起外显子2~8的重复,进而导致整个结合区的重复。对此插入片段的末端进行序列分析,显示其两端都存在*Alu*序列。*Alu*序列是大约300bp长的重复序列,含有限制酶*Alu*的识别位点,在人类基因组中存在大约90万个拷贝,主要位于内含子和基因间隔区(见第四章)。由于这些序列的高度相似性,有人提出在这些同源重复元件之间可能发生不等交换(图7-12)。*Alu*序列见于LDL受体基因的内含子1和内含子8,可能是引

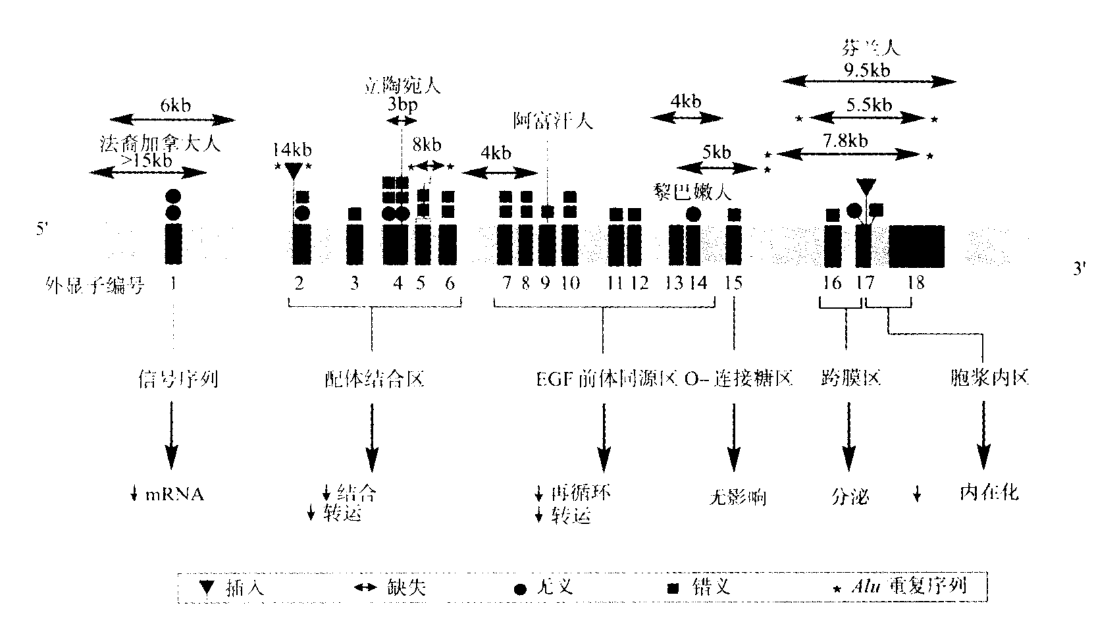


图 7-11　LDL 受体基因中选择性错义、无义、插入和缺失突变的位置。插入和缺失的大小用千碱基(kb)或碱基对(bp)表示。不同区域突变的功能后果表示在基因下面(引自 Goldstein JL, Hobbs HH, Brown MS. Familiar hypercholesterolemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995:1981～2030.)。

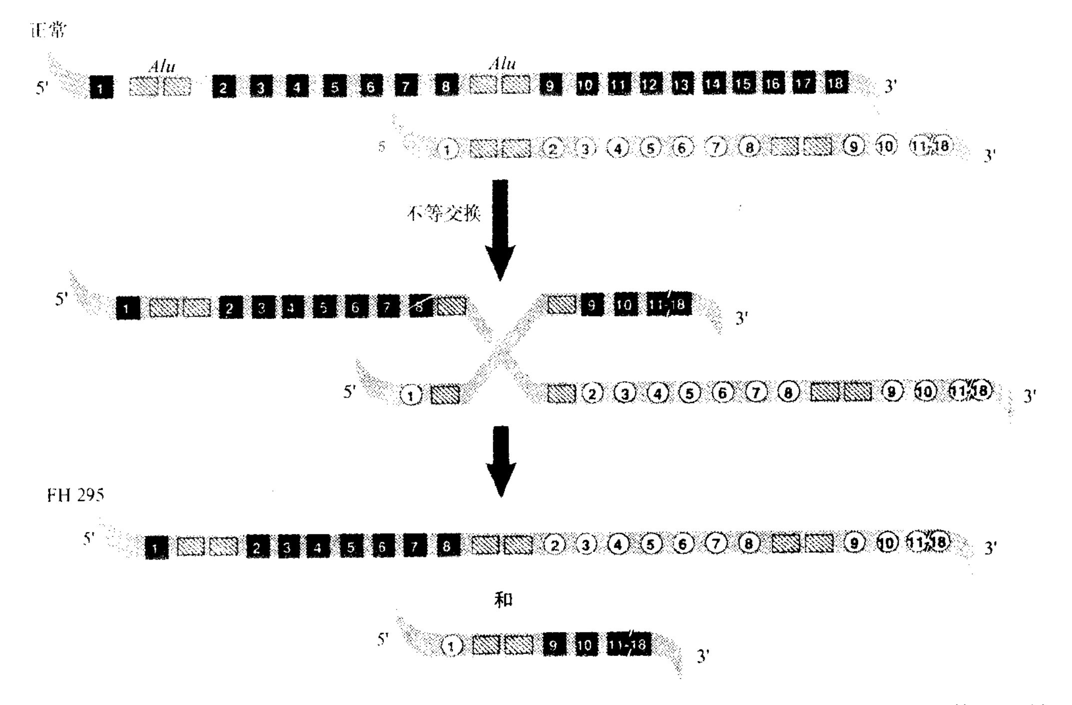


图 7-12　FH 纯合子 LDL 受体基因中两个重复 *Alu* 单元的不等交换。将两个拷贝的正常受体基因排列在一起来说明 FH295 位点的不等交换。两者的外显子分别用实心的正方形和空心的圆圈表示,交换发生于配对基因的 *Alu* 单元。重组将产生一个含两拷贝外显子 2～8 的基因和一个外显子 2～8 缺失的假拟基因。外显子用标有数字的方框表示,外显子和内含子不是按比例绘制(引自 Lehrman LE, Goldstein JL, Russell DW, Brown MS. Duplication of seven exons in the LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. Cell 1987;48:827～835. Cell Press.)。

起同源重组导致不等交换和外显子 2~8 重复的原因。可以用一个相似的机制来解释编码配体结合区一个 40 个氨基酸重复片段的外显子的进化。很可能是一个编码单个富含半胱氨酸的 40 个氨基酸肽链的祖先基因通过不等交换机制复制多次，产生了 LDL 受体分子现在的配体结合区。分析一个有 7.8kb 片段缺失的家系(缺失范围从内含子 15 中段到外显子 18，编码 3′端非翻译区域)，发现每个缺失断裂点末端都有 *Alu* 序列，表明这个缺失也是由 *Alu* 重复序列同源重组后继发不等交换引起的。到目前为止，已发现在测序的 10 个缺失断裂点中的 9 个的一端或两端有 *Alu* 序列，提示在这个或其他基因中，这可能是产生缺失或插入突变的一个共同机制(见第六章)。

各种影响 LDL 受体基因的突变可分为 5 个功能类型(图 7-13)。第一类等位基因约占总数的 20%，实际上检测不到 LDL 受体蛋白(无效等位基因)。这类突变最常见的是无义突变或移码突变，随机分布于基因的外显子中。没有截短的蛋白可能反映突变的 mRNA 或异常蛋白的不稳定性及降解迅速(如无义突变引起的某些类型的 β 地中海贫血)。第一类突变也可能发生在启动子区，这种情况下不产生 mRNA 和蛋白。第二类突变阻断新生的 LDL 受体蛋白从内质网转运到高尔基体。超过半数的 FH"纯合子"(复合型杂合子)至少有一个等位基因属于这一类型。2/3 这类突变位于配体结合区，在第 5 号

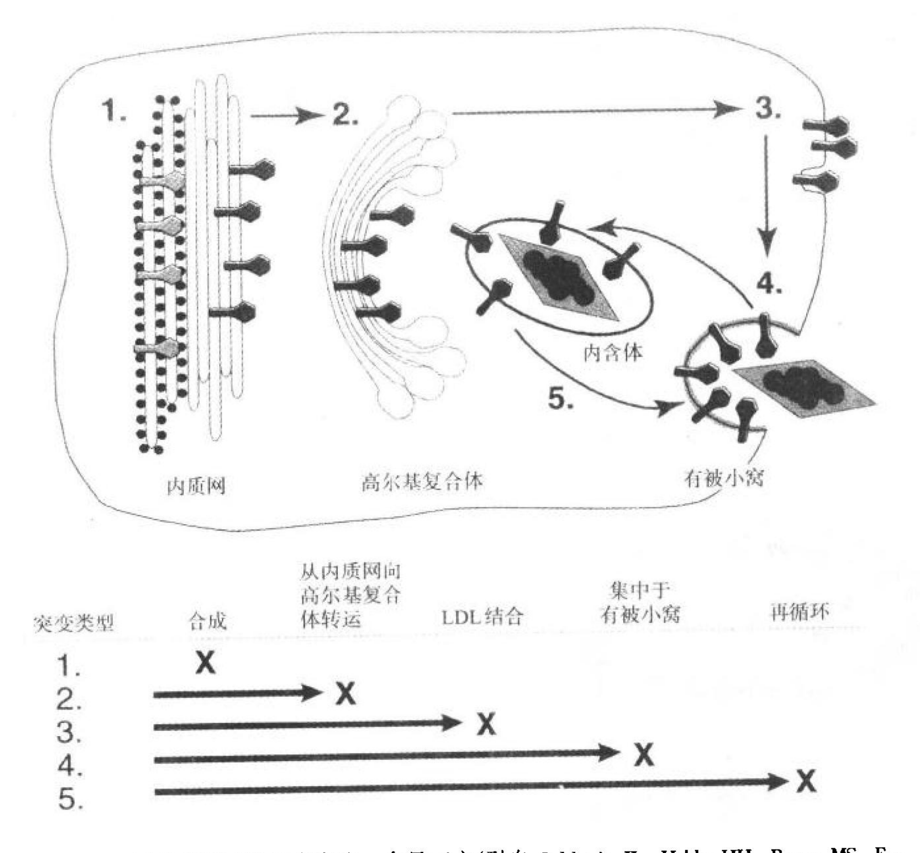


图 7-13 LDL 受体基因座的五类突变。参见正文(引自 Goldstein JL, Hobbs HH, Brown MS. Familiar hypercholesterolemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995:1981~2030.)。

重复片段有显著优势,大多为高度保守残基的错义突变。其他二类突变可见于EGF前体同源区,这是LDL受体中保守程度很高的区域。干扰新生蛋白细胞内转运的突变也是一些其他涉及膜蛋白或分泌性蛋白的遗传性疾病的一个致病原因。前面讨论过的引起α_1-抗胰蛋白酶的Pi^Z突变就是一个例子。

第三类突变编码的受体可以到达细胞的表面,但不能正常地结合配体。其原因是配体结合区或EGF前体同源区发生了符合可读框的插入或缺失。任何富含半胱氨酸的重复片段的缺失都会使LDL结合减少,因为LDL通过单个apoB-100载脂蛋白分子与受体结合,但中密度脂蛋白(IDL)与受体结合不受影响,因为IDL通过多拷贝的apoE分子与受体结合。第五重复片段的缺失是个例外,它会减少LDL和IDL与受体结合,且通常会产生一个严重得多的表型。

第四类突变编码的受体可到达细胞表面,也能正常地结合LDL,但不能集中在网格蛋白包被的陷窝,也不能内在化结合LDL。这类突变包括错义、无义、移码突变以及缺失,可能只影响受体分子的胞浆内区或跨膜区。

第一个受体错位突变发现于一个有纯合性FH临床表现的男性。他的成纤维细胞在培养时表现出接近正常的LDL结合能力,但不能内在化,所以不能调节HMG-CoA还原酶。分析他父母的成纤维细胞发现,他母亲的突变等位基因编码的是无功能的受体,而他父亲的突变等位基因编码的受体不在有被小窝内,而是随机地插入膜中,这种LDL受体能结合LDL但不能内在化。在这种突变中,807位编码酪氨酸的TAT变为编码半胱氨酸的TGT(图7-14)。通过寡核苷酸位点定向诱发突变技术使正常LDL受体cDNA突变

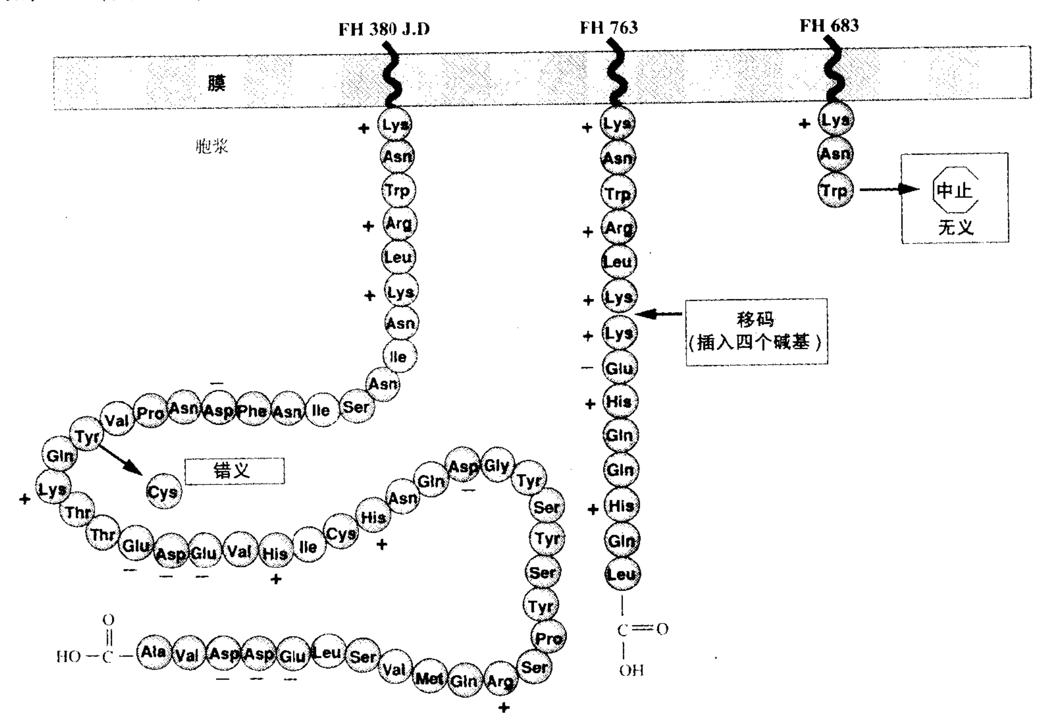


图7-14 三个内在化缺陷型FH纯合子中影响LDL受体胞浆区的突变(引自Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homerostasis. Science 1986;232:34~47. The Nobel Foundation, 1986.)。

来对患者的突变复型,证实这种突变会引起受体错位的表型。将突变的 cDNA 转染缺乏 LDL 受体的突变系中国仓鼠卵巢细胞,转染细胞在细胞膜上随机表达 LDL 受体,但在有被小窝内没有表达。其他影响 LDL 的 C 末端突变,如上面讨论过的 7.8kb 缺失,导致受体不能锚定在细胞膜上,从而使 90%以上新合成的受体分泌到细胞外。

最后,第五类突变编码的受体可结合并内在化 LDL,但不能释放内含体中的受体,也不能回到细胞表面(称为再循环缺陷型突变)。这类突变见于约 20%的 FH"纯合子",大多数是 EGF 前体同源区 5′端的错义突变,此区域内的保守区对正常受体的再循环来说是非常重要的。如果配体不能从受体释放出来,受体配体复合物将被降解。

所以,对引起 FH 的各种突变进行分析和功能分类,可以更好地理解这种重要受体的区域结构及每个结构域的功能特征。这种分析可使特定突变与某种细胞和临床表型联系起来。与突变编码完全无功能受体的患者相比,突变引起 LDL 受体功能减低但没完全失去功能的患者往往血浆 LDL 胆固醇水平较低,动脉粥样硬化性心血管疾病较轻,治疗效果也较好。虽然如此,同一家系中有同一突变类型的受累个体的 LDL 浓度和临床表型却可有较大差异。这种可复性反映了别的基因或环境因子的修饰效应。在一个有启发性的家系中,从细胞或分子水平证实 1/3 个体为 LDL 受体突变的杂合子,但其血浆 LDL 水平都正常。在此家系中,有证据表明一个与 LDL 受体基因座非连锁的常染色体显性基因减轻了 FH 突变的效应。

现在,从 DNA 水平对 FH 进行筛查和产前诊断已成为可能。对那些存在单个共同突变的群体,可直接检测突变。但是,在大多数群体(如北美人群)中,存在着多种 LDL 受体突变,如果可找到受累家系的成员,可用 DNA 多态连锁分析来进行分子诊断。如果能提供 LDL 受体基因内的多个 DNA 多态,就可构建单体型用于追踪高危家庭突变基因的分离。

(四) LDL 受体途径的生理学

尽管家族性高胆固醇血症的细胞和分子缺陷已在培养的成纤维细胞得到明确,但胆固醇生物合成和摄取的主要部位在肝脏。随后的研究表明肝脏含有与成纤维细胞相同的 LDL 受体系统,且 70%肝摄取的 LDL 是由 LDL 受体途径介导的。通过一种动物模型——Watanabe 遗传性高血脂(WHHL)兔,LDL 受体缺乏在家族性高胆固醇血症中的病理生理作用得到了进一步证实。纯合的突变兔缺乏 LDL 受体,且胆固醇显著升高,与人类家族性高胆固醇血症非常相似。由于从 IDL (它可正常被肝 LDL 受体摄取)转化的 LDL 增多,而 LDL 受体的减少又使 LDL 的清除减少,二者联合作用使 LDL 胆固醇的水平升高。对小鼠的研究进一步证实了 LDL 受体在动脉粥样硬化中的作用。通过在胚胎干细胞中进行同源重组,产生靶向破坏 LDL 受体基因的纯合小鼠,其血浆 LDL 胆固醇水平显著升高。当饲以高脂饮食时,它们发展为严重的动脉粥样硬化合并 LDL 胆固醇显著升高。野生型鼠喂同样食物则没有这些变化。最后,给受体缺乏小鼠注射编码人 LDL 受体的重组腺病毒,则可恢复其肝脏 LDL 受体活性和正常的胆固醇代谢。

（五）LDL受体模型的治疗意义

Brown 和 Goldstein 的研究预示可通过增加杂合性缺陷个体肝脏对胆固醇的需求量来增加其肝脏 LDL 受体的数目。通过减少食物中胆固醇和饱和脂肪酸的含量以降低外源性胆固醇确实可少量且有限地降低受累个体的 LDL 胆固醇浓度，而用阴离子交换树脂如消胆胺干扰胆汁酸的肠肝循环可更显著地减少血中胆固醇的含量(图 7-15)。胆汁酸含大量胆固醇，正常情况下分泌入胆汁并由回肠末端重吸收运回肝脏。通过阻碍胆汁酸的重吸收，杂合子家族性高胆固醇血症患者胆固醇水平可减低 25% ~ 30%。这可能与肝脏 LDL 受体数目增加有关。但不幸的是，肝细胞也会对经肠肝循环回收的胆固醇减少作出反应，增加 HMG-CoA 还原酶和内源性胆固醇的生物合成。内源性胆固醇的生物合成可被一类称为抑制素的药物抑制，抑制素是 HMG-CoA 还原酶强烈的竞争性抑制剂。内源性 HMG-CoA 还原酶的抑制引起 LDL 受体数目的增加，如与胆汁酸阻断剂一起使用，可使家族性高胆固醇血症的杂合子血胆固醇水平降至正常范围。在此情况下，如要保持供给肝脏正常的胆固醇量，就必须增加 LDL 受体数目，这样血中 LDL 胆固醇的浓度就会降低(图 7-15)。因此，由于弄清了 LDL 受体途径，已建立了对这种严重疾病合理而有效的治疗方法。

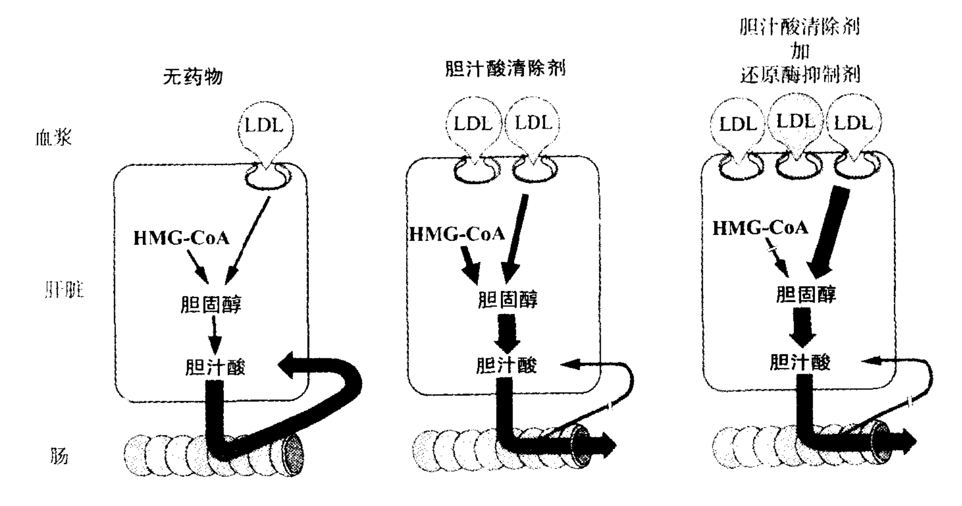


图 7-15 用胆汁酸结合树脂和 HMG-CoA 还原酶抑制剂治疗 FH 杂合子的基本原理(引自 Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homerostasis. Science 1986; 232:34 ~ 47. The Nobel Foundation, 1986.)。

对于家族性高胆固醇血症纯合子，LDL 受体数目通常不能增加。这样的患者已通过肝脏移植进行救治，通常还需同时进行心脏移植，因为由于早发动脉粥样硬化性心脏病，患者的心脏已严重受损。这种疾病的体细胞基因疗法正在积极的研究中，临床试验目前也在进行。

LDL 受体模式也许与非 FH 普通人群的高胆固醇血症和胆固醇调节也有关。对包括人类在内的几个物种的 LDL 受体的研究发现，在人类中，受体只在比一般西方人 LDL 胆

固醇低得多的浓度发挥功能。所以,在"正常"人 LDL 胆固醇浓度时受体是饱和的,且受体的合成受到抑制,降低食物中胆固醇和饱和脂肪酸的摄入也许可略增加新的 LDL 受体的合成。对非家族性高胆固醇患者用抑制素进行治疗可降低 LDL 胆固醇 25% ~ 35%。更重要的是,可减少 25% ~ 30%的心脏病的发生。

对家族性高胆固醇血症的分子病理学的研究清楚地证实了医学和遗传学之间的相互渗透。两位受过医学训练的研究者最初是研究冠心病,后将注意力集中到家族性高胆固醇血症这一先天性代谢缺陷上。他们非常机敏地将研究集中在一种罕见的、纯合的突变,即纯合的家族性高胆固醇血症上,运用从细胞生物学、生物化学、免疫学到分子生物学的一系列适当的现代技术方法,对正常和异常的胆固醇代谢的了解做出了重要的贡献。他们发现并明确了 LDL 受体的功能,建立了一种治疗家族性高胆固醇血症这种常见重症的合理而成功的疗法。沿着这个方向,他们查清并描述了受体介导胞吞的特征,这是一种真核细胞中普遍存在的重要机制。他们还对了解突变的分子机制和不等互换在基因进化中的作用做出了重要的贡献。最后,由于甲羟戊酸合成途径与各种类异戊二烯化合物的生成有关,而后者中有些与细胞生长调节有关,了解胆固醇生物合成的调节有可能导致在生长调节性疾病(如癌)治疗方面的重要启发。

五、血　友　病

据犹太法典记载,早在公元前二世纪,教士曾免除了一个男婴的割礼,因为他母亲姐姐的三个儿子都死于包皮环切术后的出血。这种对 X 连锁遗传的表面理解是对人类遗传性疾病最早的认识之一。这种出血综合征在 18 世纪晚期和 19 世纪早期被再发现是典型的血友病,且累及了好几个欧洲皇室(图 7-16)。

如图 7-17 所示,血液的凝固由一个复杂的调节通路控制,包括酶、辅助因子和抑制子。这个通路中的一些步骤如有缺陷会引起相似的出血症状。参与凝血反应的各种蛋白很容易被研究人员获得,由于这个原因,这些蛋白在医学发展早期是人们着力研究的目标。此外,几乎所有这些组分都存在缺陷的病人,这对凝血研究的进展也非常重要。临床上最重要的遗传性出血性疾病是甲型血友病,由凝血因子Ⅷ缺乏引起。Ⅷ因子是丝氨酸蛋白酶Ⅸ因子的一种非酶的辅助因子。Ⅷ因子或Ⅸ因子缺乏会导致相似的正常凝血过程的阻断,引起临床上难以区分的出血。Ⅸ因子缺乏称为乙型血友病。由于Ⅷ因子和Ⅸ因子都位于 X 染色体上,甲型和乙型血友病都是 X 连锁隐性遗传病。

甲型血友病人的严重程度很不一致,与其血中存留的Ⅷ因子活力直接相关。与许多由酶功能异常引起的疾病一样,Ⅷ因子水平低于正常 5%开始出现轻微的症状,高于 25%一般表现为凝血功能正常。Ⅷ因子完全缺失(< 正常 1%)属于严重的甲型血友病,可见于 70%的病例。重症血友病患者频发出血,常为自发性或微小创伤引起,关节腔内出血特别常见(图 7-18)。关节腔内反复出血会引起慢性损伤,常导致严重的功能障碍,有时需关节移植。出血也可发生于许多其他器官,包括大脑,这是患者常见的死亡原因。血友病的治疗采取补给缺乏的因子,最早是输全血,后来使用纯化的浓缩Ⅷ因子。这种Ⅷ因子浓缩制剂是用成千上万的献血者的血液制备的,不幸的是,在20世纪80年代中期以前

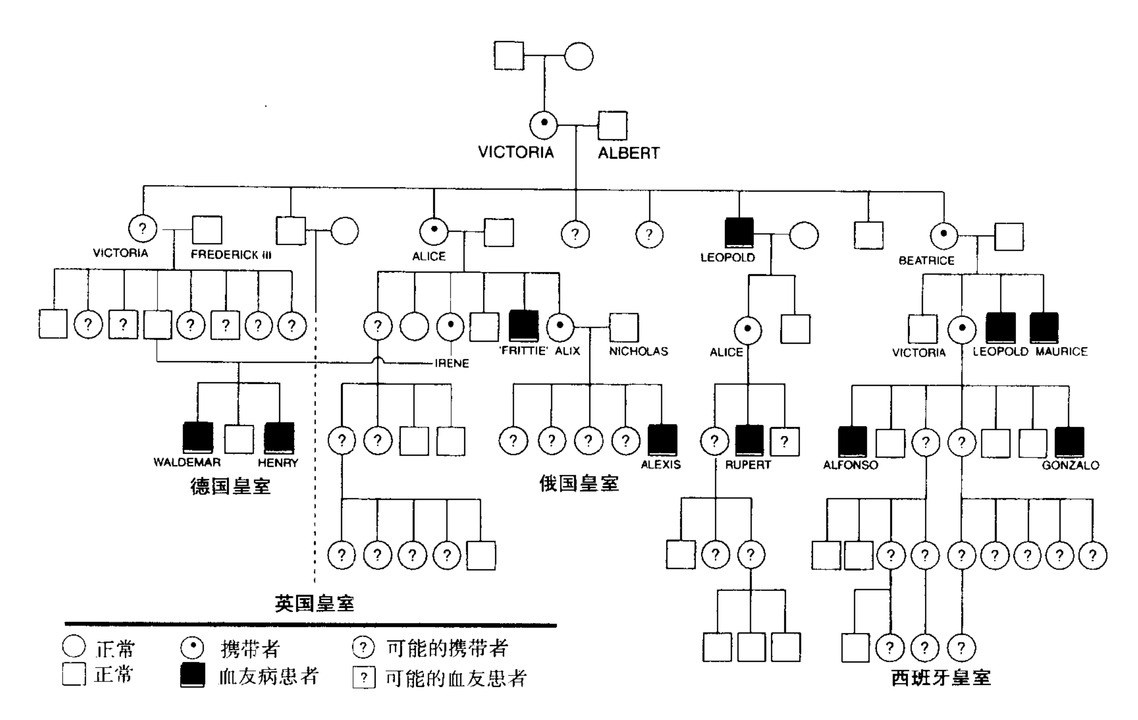


图 7-16　Victoria 女王的系谱图。尽管 X 连锁隐性遗传方式说明缺陷可能是Ⅷ因子基因或Ⅸ因子基因，但当时并不知道这两种疾病（甲型血友病和乙型血友病）的区别。Victoria 女王是甲型血友病携带者，将血友病等位基因传递给了她受累的儿子 Leopold 及女儿 Alice 和 Beatrice，这两个女儿都有儿孙受累。现在的英国皇室血统来自一个正常男性，所以未受累（引自 Kazazian HH, Tuddenham EGD, Antonarakis SE. Hemophilia A and parahemophilia. Deficiencies of coagulation factors Ⅷ and Ⅴ. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995:3241～3267.）。

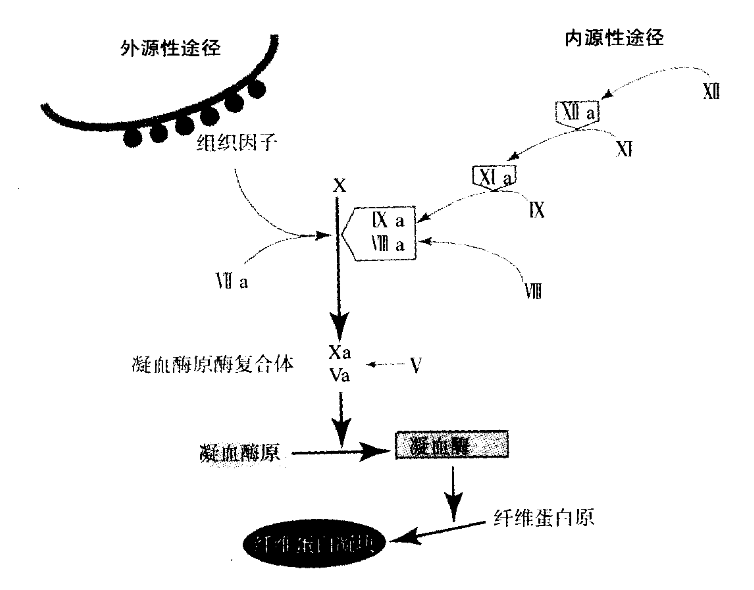


图 7-17　血液凝固的级联式过程。现在一般认为凝血过程主要通过组织因子启动，内源性途径中包括Ⅷ因子和Ⅸ因子，它们的主要功能是一扩增环，复合体将Ⅹ因子活化为Xa，形成凝血酶原复合物，进而生成凝血酶，导致纤维蛋白血凝块沉淀。缺乏Ⅷ因子或Ⅸ因子都会引起 X 因子活化的减少，继发凝块形成缺陷，从而出现血友病的出血表型。

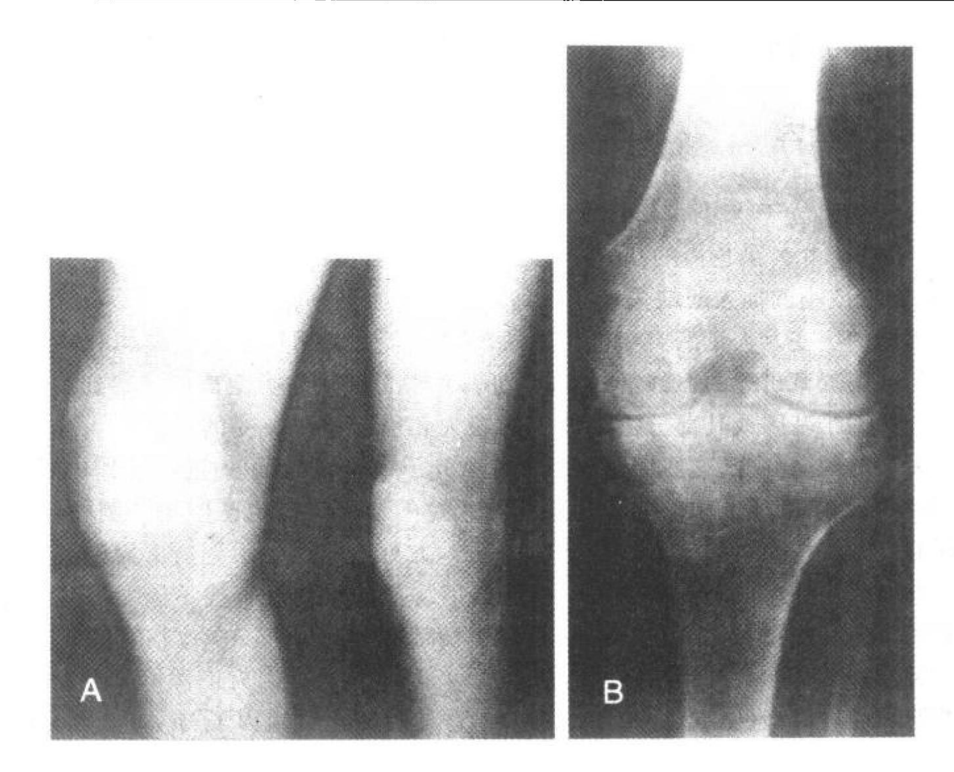


图 7-18 甲型血友病的关节积血。(A)示右膝急性关节出血。(B)为一 X 光片,示由于反复关节积血引起的血友病性关节病的慢性破坏性骨改变:骨变薄,踝关节增大,软骨缺失(引自 Ginsburg D. Hemophilia and other inherited disorders of hemostasis and thrombosis. In: Rimoin DL, Connor JM, Pyeritz RE, Emery AEH. Emrey and Rimoin's principles and practice of medical genetics. 3rd edition New York: Churchill Livingstone, 1996.)。

使用的方法没有灭活常见的病毒污染物。结果,在 1979 年 ~ 1985 年之间使用Ⅷ因子浓缩品的患者约 90%因疏忽而感染了 HIV 病毒,他们中的许多人死于艾滋病,而艾滋病治疗也成为血友病临床治疗的一个不可分的部分。由于对献血者进行 HIV 筛查以及Ⅷ因子制备方法的改进,1985 年以后开始治疗的病人有效地排除了感染 HIV 的危险。1984 年,两家生物技术公司克隆了Ⅷ因子基因,这在很大程度上是因为想得到一种安全的重组Ⅷ因子产品,这种产品最后于 1994 年批准供患者使用。分析Ⅷ因子基因也使得鉴定无数种甲型血友病突变型成为可能,下面将讨论这一点。

甲型血友病在男性中的发病率尽管经常被说成是 1/10 000,其实更可能高达1/5 000,所以突变型等位基因频率(q)也是 1/5 000。甲型血友病杂合子女性携带者(频率为 1/2 500)大约有 50%的正常Ⅷ因子水平,完全没有症状。女性血友病人由于是突变等位基因的纯合子,其临床发病率预测大约为 1/25 000 000(1/5 000 × 1/5 000, q^2)。女性甲型血友病也可能是由于 X 染色体极度扭曲或易位或其他导致 X 染色体失活的 X 染色体缺陷引起的(见第八章)。1935 年,Haldane 提出,患有致死的 X 连锁疾病的男性中 1/3 是新的突变。这个预言,有时也称作 Haldane 假说(见第四章),根据这个假设提出:1/3 X 染色体上的突变型等位基因由男性携带,也将会在每一代人中丢失。所以,如果人口与疾病发生率保持平衡,这些等位基因必须由等量的新突变来替代。对Ⅷ因子基因突变的研究事

实上已经证实了这个假设。

(一) Ⅷ因子基因

Ⅷ因子基因位于X染色体长臂末端(Xq28)(图7-19),是一个大基因,含26个外显子,全长186kb,占X染色体的0.1%。到目前为止,已鉴定出Ⅷ因子基因78种大的缺失、223种点突变或小的插入或缺失。考虑到新突变发生频率高,通常认为无亲缘关系的家系携带不同的突变。大量的突变产生的缺陷提供了有关Ⅷ因子蛋白结构和功能的大量信息。如同第五章讲述的血红蛋白病的范例,几乎每种缺陷类型都已观察到。它们包括引起整个Ⅷ因子基因或其部分丢失的大的缺失和引起移码突变的小的插入或缺失。已鉴定出无数的无义突变,这些突变分布于整个分子中,导致截短的蛋白及其功能的丧失。最后,影响Ⅷ因子活化或功能的特定步骤的大量的错义突变也已鉴定。例如,发现了多种不同的氨基酸替换代替372密码子的精氨酸,此处是Ⅷ因子活化所必需的凝血酶切割位点。丧失活化位点使血友病表型至少部分因特异的氨基酸替换不同而发生从轻度到重度的变化。有些突变也发生于另一个凝血酶活化位点,即1689位精氨酸。1680位的酪氨酸突变为脯氨酸会产生中度严重的血友病表型,可能是因为正常情况下发生于这个酪氨酸的翻译后修饰(硫酸化)功能的丧失。这种硫酸化对Ⅷ因子与von Willebrand因子的相互作用是必需的。

大约1/3 Ⅷ因子的点突变发生在CpG双核苷酸这一人类基因组的突变"热点"。尽管双核苷酸也许只有1/16的可能是CpG,但约1/3的Ⅷ因子点突变发生在这些位点。这类位点的高频率突变也见于其他基因,如本章前面讨论过的G6PD。

由于存在大量不同可能的突变,且Ⅷ因子基因很大,一般很难在一个特定的家系中鉴定出特异的缺陷。如第五章介绍的那样,如果能找到多个受累的家系成员,就可对Ⅷ因子基因内的DNA序列多态性进行遗传连锁分析,从而进行产前诊断和携带者的检查。但是,在仅有一个受累成员的家系中,连锁分析不可行。现在已建立了很多筛查方法(第五章已论及)来检测特异的突变。用这些方法筛查全部外显子和内含子/外显子边界,可对几乎所有血浆中有异常Ⅷ因子蛋白的患者进行突变的检测。毫不奇怪,这些患者中几乎都可找到编码序列的错义突变,导致单个氨基酸替换,从而引起Ⅷ因子功能的丧失或减低。

(二) 重复发生突变的新机制

然而,用上述筛查方法,在无可检测Ⅷ因子蛋白的患者中只有一半可鉴定出突变,还有大约5%的患者可通过DNA印迹分析检测到较大的缺失,剩下的40%~50%重症血友病患者检测不到缺陷。现在发现几乎所有这样的患者都存在同一种缺陷,即X染色体远端的一个片段倒位,致使Ⅷ因子基因破坏。图7-19说明了这种倒位机制。这个倒位是Ⅷ因子基因内含子22中一个称为A基因的小基因与Ⅷ因子上游朝向X染色体末端的两个A基因额外拷贝中的一个发生同源重组。产生的这种倒位由于切掉了23~26外

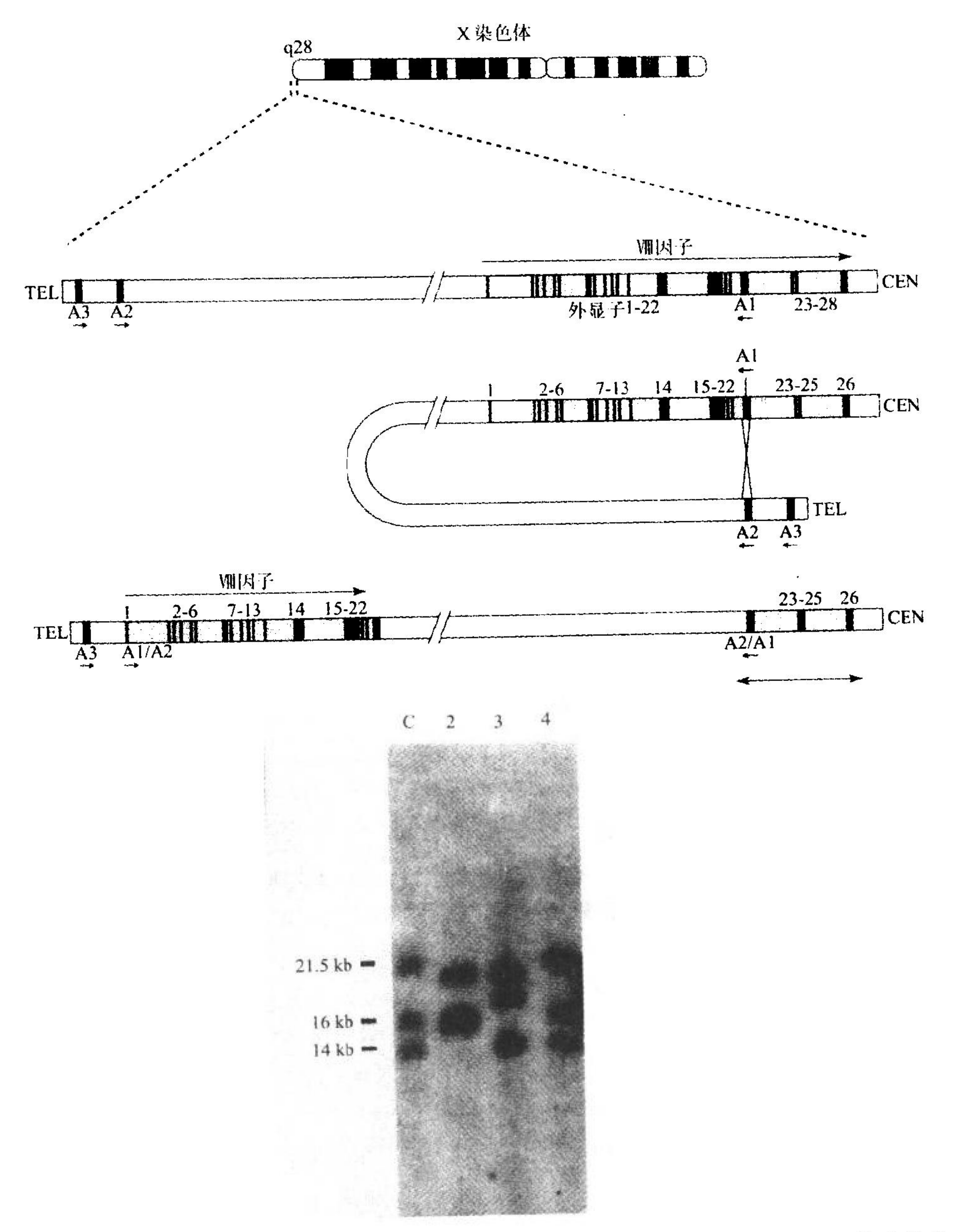


图 7-19 Ⅷ因子基因模式图及共同的血友病基因倒位机制。Ⅷ因子基因位于 X 染色体长臂末端附近。A 基因在 X 染色体的这个区域内存在三个拷贝，其功能尚不清楚。其中一个位于Ⅷ因子基因的第 22 内含子内(A1)，另两个靠近 X 染色体末端(A2、A3)。A1 基因和两个上游拷贝的一个(A2 或 A3)发生同源重组，引起一个含 1~22 外显子的Ⅷ因子基因片段倒位，这种倒位导致Ⅷ因子功能完全丧失和严重的甲型血友病。这里将重组表示为 A1 和 A2 之间进行，但 A3 也可发生。图下方的 Southern 印迹表示患者的 DNA 经限制酶 *Bcl* Ⅰ消化后，以 A 基因作为探针进行杂交的结果。C 道为一对照 DNA 样品，21.5kb 带为 A1，14kb 带为 A2，16kb 带为 A3。2 道为 A1 与 A2 发生重组倒位的血友病患者 DNA，3 道示 A1 与 A3 之间的重组，4 道为未发生倒位的血友病患者 DNA。两种类型的倒位 A1 特异带的大小都有改变，因为带中还有重组的另一 A 基因拷贝(A2 或 A3)(引自 Lakich D, Kazazian HH Jr. Antonarakis SE, Gitschier J. Inversions disrupting the factor Ⅷ gene are a common cause of severe haemophilia A. Nat Genet 1993;5:236~241.)。

显子而破坏了编码蛋白 C 末端的Ⅷ因子基因。值得注意的是，这种相同的突变在许多无亲缘关系的病人中重复发生。这种独特的分子机制提供了在大量患者中进行 DNA 诊断的方法。DNA 印迹分析检测这种倒位，可以鉴定大约 45% 重症甲型血友病患者的缺陷。

这种独特的染色体内同源重组在其他人类疾病中没有先例。最近的研究表明这种现象只发生在男性减数分裂过程，可能是因为需要一个未配对的 X 染色体。这种发现有重要的临床应用价值，有这种缺陷的儿童的母亲通常可推测为携带者，即使没有家族史，通过重组现象也可追溯到其外祖父。

(三) 乙型血友病

乙型血友病由Ⅸ因子缺乏引起，发病率约为甲型血友病的 1/10。Ⅸ因子基因与Ⅷ因子基因相距约 10cM，长 34kb，含 8 个外显子。与甲型血友病和Ⅷ因子相似，乙型血友病与大量的Ⅸ因子突变有关。点突变可发生在整个编码序列，其中 1/3 位于 CpG 双核苷酸。同样，如 Haldane 假说所预言的那样，有 1/3 的突变似是新发生的。然而，与甲型血友病不同，没有发现乙型血友病存在共同的重复突变机制，这使常规的突变检测很困难。一个称为 Leiden Ⅸ因子的独特的乙型血友病变异型有一定的启发性。这种病人患有严重的血友病，但在青春期可显著改善。研究发现患者的Ⅸ因子基因启动子发生了突变，使得基因需要雄激素的刺激才能高水平表达。所以，其Ⅸ因子水平在儿童期很低，但在青春期会增加。

图 7-17 所示的一些其他凝血因子的缺陷可引起与甲型和乙型血友病患者相似的表型，但是，由于相应的基因都位于常染色体上(另一个基因座异质性的例子)，遗传方式为常染色体隐性遗传。这些疾病大多数很罕见。

最近的分子遗传学进展对遗传性疾病产生了重要的影响，血友病就是一个例子。突变的鉴定为甲型血友病的 DNA 检测(包括产前诊断)提供了有力的工具。Ⅷ因子基因的克隆使重组Ⅷ因子得以生产，这为许多患者提供了有效而安全的替代疗法。最后，血友病还是用体细胞基因疗法进行长期治疗的一个很具吸引力的目标。如输入Ⅷ因子浓缩品所证实的那样，仅将缺失的蛋白输入血浆就可完全消除出血症状。此外，基因表达的精确调控这个当前基因治疗策略难以克服的障碍在此并无必要，因为即使蛋白水平的轻微增加也可引起临床症状的显著改善，而增加几倍也不会引起任何明显的并发症。早期实验动物的研究证实了通过体细胞基因疗法替换Ⅷ因子和Ⅸ因子的可行性，可以想像，这种方法将来会成为甲型血友病治疗标准的一部分。

六、胶原及其疾病

先天性代谢缺陷是由某种酶的基因发生突变引起，主要以隐性方式遗传。正如常染色体隐性遗传病杂合子那样，大多数酶减少到 50% 的量仍然够用，不会引起任何表型异常。但是，在家族性高胆固醇血症的例子中，突变发生在细胞表面受体基因，受体减少到

表 7-2　纤维状胶原的主要类型

胶原类型	链	胶原基因	染色体定位	原胶原分子	组织分布	疾　病
Ⅰ	α1(Ⅰ)	*COL1A1*	17	$[\alpha1(Ⅰ)]_2\alpha2(Ⅰ)$	骨、腱、皮肤	成骨不全
	α2(Ⅰ)	*COL1A2*	7			
Ⅱ	α1(Ⅱ)	*COL2A1*	12	$[\alpha1(Ⅰ)]_3$	软骨、玻璃体	软骨发育不全
Ⅲ	α1(Ⅲ)	*COL3A1*	2	$[\alpha1(Ⅲ)]_3$	皮肤、动脉、子宫、肠	EDS Ⅳ

28 个胶原基因分布于整个人类基因组中，位于 12 条不同的染色体上。这些基因很大，长约 40kb，含有多达 52 个外显子。

胶原的中心三聚体螺旋区是一种很特别的蛋白结构。在此区域，每种胶原的每条肽链都由重复的氨基酸序列组成，可以表示为 $(Gly\text{-}X\text{-}Y)_n$。在纤维胶原（Ⅰ、Ⅱ、Ⅲ、Ⅴ和Ⅸ型）中，有一个包括 300 多个这种三联体的完整序列。X 位置常为脯氨酸，Y 位置常为稀有氨基酸羟脯氨酸或羟赖氨酸，后两者是脯氨酸和赖氨酸在翻译后修饰中产生的。三聚体螺旋位于一个限制性区域，三联体的每第三个氨基酸位于其中央，只能容纳甘氨酸这个最小的氨基酸，X 和 Y 位置氨基酸的侧链位于三聚体螺旋的表面。氨基酸三联体重复对胶原纤维的正常组装和分子的稳定性都很重要。突变，特别是甘氨酸的突变，容易引起三联体螺旋的不稳定，使得胶原纤维不能形成正常的构象和长度。

（二）遗传性胶原病

两种分类属于“遗传性结缔组织病”的疾病，即成骨不全和 Ehlers-Danlos 综合征，其分子遗传学已于最近得以明确。对其基本发病机理的认识阐明了疾病的一种新的分子机制，即“蛋白自杀”。

1. 成骨不全

成骨不全（OI）一词常用于描述一组以骨骼易碎为特征的异质性临床疾病。Ⅰ型 OI 最常见，大约有 1/15 000 的个体受累，以常染色体显性方式遗传。受累个体从出生起巩膜即有明显的蓝色，从婴儿期起易频繁发生骨折，但青春期往往减少。其上肢和下肢的长骨、肋骨、手和足的小骨最常骨折，但可正常愈合且通常不会畸形。在约半数的有Ⅰ型 OI 的家系中，受累个体在青少年晚期出现听力丧失，可能是中耳内骨骼硬化的结果。

对培养的Ⅰ型 OI 患者的成纤维细胞进行生化分析显示，这些细胞只合成正常量一半的Ⅰ型原胶原的前 α1(Ⅰ)链。在分子水平已发现各种突变，包括多个外显子的缺失或插入及无义突变，这使功能性前 α1(Ⅰ)链的合成只有正常量的一半，在培养的成纤维细胞的胞浆内几乎没有突变等位基因的产物。由于Ⅰ型胶原形成正常的三聚体螺旋所需 α1(Ⅰ)链和 α2(Ⅰ)链的比例为 2:1，过剩的 α2 链会降解掉。所以，如图 7-21 所示，只有一半正常量的Ⅰ型胶原生成。这大概就是这种疾病以骨骼脆性为特征的原因。所以，不同于蛋白产物为酶的情况，结构蛋白的量减少到 50% 会产生明显的表型后果。

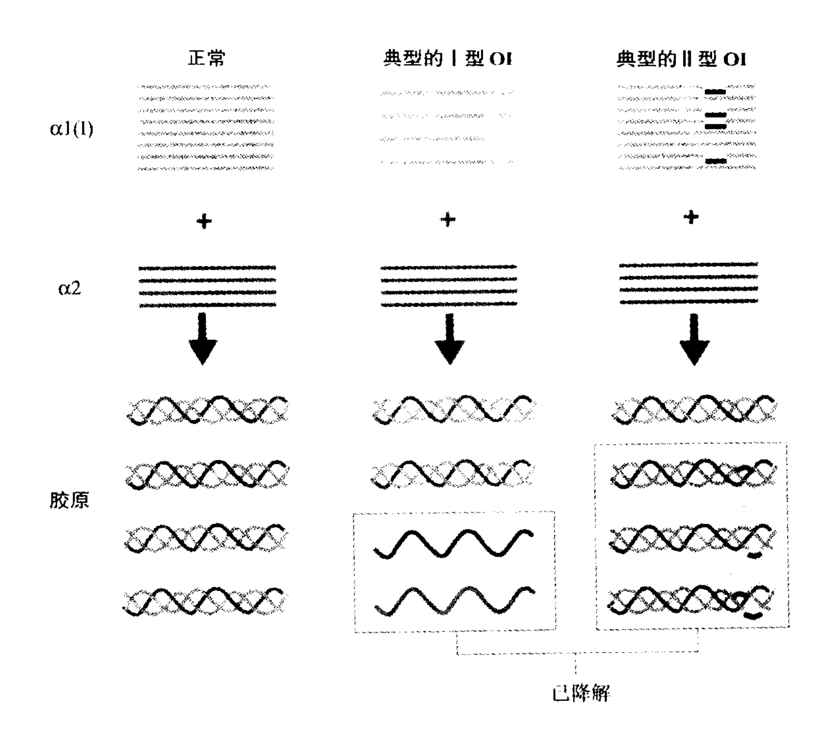


图 7-21　Ⅰ型和Ⅱ型成骨不全中胶原缺陷的机制。这两种疾病的分子原因是异质性的。这里显示的是典型的例子。在Ⅰ型 OI 中，一条 α1(Ⅰ)等位基因失活使得这种多肽链只有正常量的 50%，原胶原分子的产量也减少 50%，多余的两条 α2 链被降解掉。而在Ⅱ型 OI 中，典型的情况是一条 α 链的质量异常，如这个例子，其 α1(Ⅰ)链中的氨基酸替代用黑色表示。如果这种替代抑制正常三聚体螺旋的形成，那么 3/4 的三聚体都是异常的且将被降解，只有 1/4 幸免。

Ⅱ型 OI 是一种严重得多的疾病，患者出生时即有数不清的骨折，常在出生后几周或几个月内死亡。图 7-22 显示一名患者的 X 光片，表明存在大量的先天性骨折和长骨的缩短的原因是由于支持物的大量减少。这里胶原缺乏非常明显。由于矿化的骨骼缺乏蛋白支架，而后者的作用类似于钢筋混凝土中的钢条，所以胎儿在子宫内的常规活动也很容易引起骨折。Ⅰ型 OI 较轻，其突变常为无效突变，可使编码 α1(Ⅰ)胶原的一条 *COL1A1* 等位基因完全失活，而Ⅱ型 OI 的突变则细微得多。典型的是影响甘氨酸残基的点突变或 α1(Ⅰ)或 α2(Ⅰ)基因(*COL1A1* 和 *COL1A2*)的重排，受累的等位基因产生的异常蛋白量相对正常。这样细小的突变引起如此严重的疾病，这一事实显示出一个很具挑战性的奇怪现象，但目前已得到解决。突变型胶原肽的产量正常或接近正常，仍然可以掺入新生的原胶原三聚体螺旋分子。如图 7-21 所示，突变型 α1(Ⅰ)基因影响一半的 α1(Ⅰ)肽链，但 3/4 的胶原三聚体螺旋将是异常的。这些异常的三聚体难以组装，也许是难以运出细胞，还可能改变纤维结构，或者很快降解，总的结果是正常胶原的严重缺乏。所以，*COL1A1* 基因蛋白产物性质上的异常比数量上的异常对表型的影响严重得多。这种机制被恰当地命名为"蛋白自杀"，对于突变影响多聚体蛋白中一个亚单位所造成的严重表型，它提供了一种通用的解释，如果整体中包括一个异常的亚基，它常常导致整个多聚

体功能的丧失或减低。

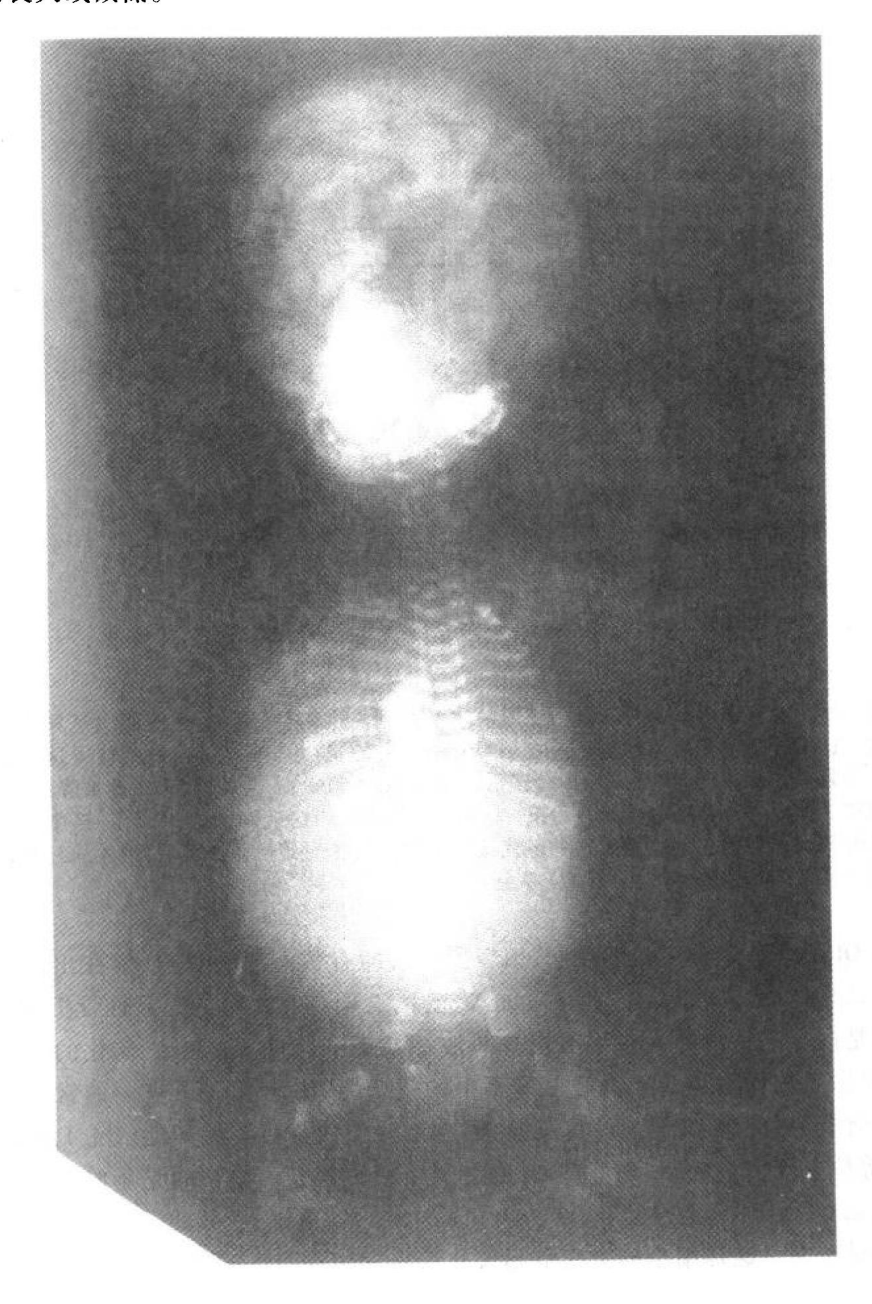


图 7-22　一名产期致死型成骨不全(Ⅱ型 OI)婴儿的 X 光片。注意压缩的股骨有多处骨折。

Ⅱ型 OI 是一种不能生育的致死性疾病。因为绝大多数病例是散发,以前一直被认为是常染色体隐性遗传病。但是,分子分析显示受累个体通常是 *COL1A1* 或 *COL1A2* 发生新突变的杂合子,后者较少见。由于新突变可能发生在胚胎发育早期或配子阶段,Ⅱ型 OI 患儿的父母中存在生殖细胞嵌合体(见第三章),这样的夫妇有 5%～10%的风险第二

个孩子也患Ⅱ型OI。如果一个患儿的分子缺陷已被证实,就可利用DNA分析进行产前诊断,通过超声波检查也可发现胎儿的骨骼异常。

2. Ehlers-Danlos 综合征

像成骨不全一样,Ehlers-Danlos综合征(EDS)也被分为很多临床亚型。这类疾病的特点是组织弹性和脆性增加,以关节松弛、皮肤变薄并产生异常疤痕为主要特征,其他特征在不同类型间彼此变化差异较大(表7-3)。图7-23示一位Ⅰ型EDS患者。这是一种常染色体显性遗传病,以皮肤和关节的过度伸展为特征,患者皮肤很薄且难以愈合,在受到相对较小的创伤后留下萎缩的"烟锡纸"样疤痕。然而,大血管的弹性总的来说未受损害。Ⅰ型EDS的分子缺陷可能是编码Ⅴ型胶原的α1(Ⅴ)链的*COL5A*基因发生了突变。



图7-23 Ⅰ型EDS患者的典型外貌。注意过度松弛的皮肤和小腿上的烟锡纸样疤痕(引自McKusick VA. Heritable disorders of connective tissues. 4th ed. St Louis: Mosby, 1972)。

表 7-3　Ehlers-Danlos 综合征[a]

类　型	临床特征	遗传方式[b]	缺　陷	突变基因
Ⅰ,Gravis	皮肤过度伸展,关节过度伸展,烟锡纸样疤痕	AD	Ⅴ型胶原异常	*COL5A*
Ⅱ,Mitis	与Ⅰ型类似但较轻	AD	未知	
Ⅲ,家族性过度伸展	显著的关节过度伸展	AD	未知	
Ⅳ,动脉型	薄而透明的皮肤,明显的青肿,动脉及肠破裂	AD	Ⅲ型胶原异常	*COL3A1*
Ⅵ,眼型	皮肤变软,眼组织变脆	AR	赖氨酸羟化酶缺乏	*PLOD*
Ⅸ,枕骨角综合征	皮肤松弛,角状枕骨,膀胱憩室	XLR	铜利用异常	*MNK*

a. 只列出一部分,最少已报道 10 种 Ehlers-Danlos 综合征。
b. AD,常染色体显性遗传;AR,常染色体隐性遗传;XLR,X 连锁隐性遗传。

了解得最清楚和最危险的 EDS 类型是Ⅳ型,其皮肤弹性和关节的活动性小一些,特征是皮肤很薄几乎为半透明,且容易受伤。但最严重的后果是由于其大血管的脆性显著增加,患者可因主动脉、脾动脉或肾动脉的破裂而突然死亡。自发性的乙状结肠穿孔也很常见。手术修补这些大的损伤一般极其困难,因为支持组织很脆,经受不住缝合。

Ⅳ型 EDS 以常染色体显性方式遗传,几乎所有受累的家系都表现出Ⅲ型胶原在质量或数量上的异常。Ⅲ型胶原是 α1(Ⅲ)链的同源三聚体(注意避免混淆表示 EDS 类型的罗马数字和代表胶原亚型的罗马数字)。在大多数情况下,质量异常出现在 *COL3A1* 基因的产物,通常源于甘氨酸密码子的错义突变或外显子跳跃突变。这样的突变引起的结构异常占三聚体螺旋异常的 7/8(通过图 7-22 阐明的蛋白自杀机制)。所以,像上面讨论过的 OI 一样,质的突变可能比无效突变引起的胶原产物异常更严重。由于胶原的生物合成缺陷已在培养的成纤维细胞表达,且特异的突变也在一些病人中得到明确,现在有可能在症状出现前甚至产前对受累家系的成员作出诊断。

其他类型的 EDS 中,另有两种分子机制不属于胶原基因突变,这里应该提及。在Ⅵ型 EDS 中,已发现了赖氨酸羟化酶的缺乏,翻译后赖氨酸的羟化需要此酶。赖氨酸位于上面提到的 Gly-X-Y 重复序列的 Y 位置,没有这种羟化作用,三聚体螺旋的交联和稳定性都会减低,而且胶原纤维会缺乏必要的强度。Ⅸ型 EDS 中已证实缺乏另一种酶,即赖氨酸氧化酶。这种酶需要铜作辅助因子,氧化赖氨酸和羟基赖氨酸,使胶原进一步交联。Ⅸ型 EDS 是一种 X 连锁隐性遗传病,现在已知道其分子缺陷位于 Menkes 综合征基因,后者与铜代谢有关。所以,EDS 同时表现出等位基因异质性(EDS Ⅳ中 *COL3A1* 基因发生不同的突变)和基因座异质性(不同的基因座突变引起 EDSⅠ、EDSⅥ、EDSⅣ和 EDSⅨ)。

3. 其他胶原疾病

Ⅱ型胶原维持软骨的强度,其突变会引发一组称为软骨发育不良的遗传性疾病,以

矮小、关节畸形和其他骨骼异常为特征。像成骨不全的情况一样，这些疾病是由Ⅱ型胶原基因（*COL2A1*）发生各种不同的突变包括缺失、插入或甘氨酸替换引起的。

（三）从胶原疾病得到的启示

我们讨论这类疾病的目的一方面是为了介绍一组重要的人类遗传病，另一方面也是为了就疾病的分子基础得出一些普遍的结论。第一，显性遗传病的蛋白自杀机制，借此一个突变的亚单位常引致整个多亚基蛋白的丢失，首先是在分析胶原疾病时搞清楚的，随后在其他疾病中也有发现；第二，按临床标准将疾病分类，即使规定得很详细，也经常低估遗传异质性的程度，这只有在疾病的分子基础明确后才能被认识。这在胶原疾病中特别明显，因为涉及一些基因座、基因很大、且潜在的突变数目几乎是无限的。这种高度遗传异质性的自然结果是一些引起中性表型的突变不能随便归入任何临床类型；第三，即使获得胶原蛋白相当多的信息，预测基于突变的临床表型仍然不是完全可能的，所以，至少在目前，准确预测严重程度通常只能参考受累家系其他成员。最后，这些疾病证实表型的变异很大（成骨不全是一个很好的例子），提示诸如骨质疏松症（"骨骼变薄"）和骨关节炎（退化性关节病）等成人很常见疾病表型不一致可能是由于其发病基础在于胶原基因的微小改变。实际上，现在有一些证据表明 *COL1A2* 突变可能与骨质疏松症有关，而 *COL2A1* 突变也在一个很大的骨关节炎家系中被发现。而且，还在主动脉瘤家系中发现了 *COL3A1* 的突变。

推 荐 读 物

先天性代谢性缺陷

Garrod AE. The Croonian lectures on inborn errors of metabolism. Lancet 1908;2:1-7,73-79,142-148,214-200. *The classic publication defining the concept of the inborn errors of metabolism.*

Harris H. The principles of human biochemical genetics. 3rd ed. New York: Elsevier Scientific, 1980. *A beautifully written exposition of the subject in the "premolecular" era.*

Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995. *Encyclopedic and authoritative; the standard reference on inborn errors of metabolism and inherited metabolic disease.*

尿黑酸尿症

LaDu BN. Alkaptonuria. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995: 1371 ~ 1386.

葡萄糖-6-磷酸酶缺乏症

Luzzatto L, Mehta A. Glucose-6-phosphat dehydrogenase deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995:3367 ~ 3398.

α_1-抗胰蛋白酶缺乏症

Cox DW. α_1 - Antitrypsin deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995: 4125 ~ 4158.

家族性高胆固醇血症

Goldstein Jl, Hobbs HH, Brown MS. Familiar hypercholesterolemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995: 1981 ~ 2030

Motulsky AG. The 1985 Nobel Prize in physiology or medicine. Science 1986;231:126 ~ 129. *An excellent perspective on the importance of the familial hypercholesterolemia story to both genetics and medicine.*

Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. Science 1986;232: 34 ~ 37. *The Nobel Prize acceptance lecture; an elegant, detailed account of the molecular pathology and genetics of familial hypercholesterolemia.*

血友病

Hoyer LW. Hemophilia A. N Engl Med 1994;330:38 ~ 47.

胶原疾病

Byers PH. Disorders of collagen biosynthesis and structure. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. 7th ed. New York: McGraw-Hill, 1995:4029 ~ 4077.

Beighton P. ed. McKusick's heritable disorders of connective tissue. 5th ed. St. Louis: Mosby, 1993.

Prockop DJ. Mutations in collagen genes as a cause of connective tissue disease. N Engl J Med 1992;326:540 ~ 545.

第八章　细胞遗传学

难以置信的是染色质微粒或其他任一物质无论怎样复杂，都具有必定要归因于我们的因子(基因)的多种能力。在任何已知实验中几乎都是均质的、互相难以区分的染色质微粒能按其物质本质赋予生命全部的特征，这一假说甚至超出了最令人信服的唯物论的范畴。

——William Bateson，1916（引自 Borek E. The Code of Life. New York：Columbia University Press，1964）

至少有一半的自发性流产是由染色体异常引起的，同时染色体异常也是导致先天性畸形的重要原因。0.5%以上的新生儿出生时都伴有明显的常染色体或性染色体异常。在这些染色体异常中，最常见的、也是最为人们所熟知的染色体疾病是 21 三体性(trisomy 21)，即 Down 综合征，和脆性 X 染色体综合征。除了先天性或来源于生殖系的染色体异常，发生在体细胞内的染色体异常在肿瘤的形成过程中起重要作用。我们在第二章时讲述过染色体结构及其行为。在这一章内，我们将更为详细地讨论这些题目和一些更常见的、影响染色体的异常。

直到 20 世纪 50 年代，有关人类染色体的两个显著事实才得到确认。然而，它们两个都是错误的。第一，人们认为人类细胞含有 48 条染色体；第二，哺乳动物的性别决定被认为和果蝇(*Drosophila*)一样，是由常染色体和性染色体数目比率决定的。当时只能在分裂活跃的细胞如肿瘤细胞，男性性腺细胞、组织培养中的生长细胞和骨髓标本中可以观察到染色体。那时染色体标本的质量与现在的相差很远(或许归因于 Bateson 的怀疑主义)。在 1956 年，Tijo 和 Levan 明确证实培养后的胚胎成纤维细胞含有 46 条染色体而不是 48 条染色体。

在三年内，三种主要的由染色体数目改变而造成的疾病被确定，它们是 Down 综合征，Turner 综合征和 Klinefelter 综合征。稍后，我们将在本章对这些疾病进行讨论。

一、细胞遗传方法学

促进细胞遗传学领域快速发展的主要进展是在外周血淋巴细胞中检测染色体技术的发展。这项技术现已常规用于检查人类染色体，即在组织培养基中加入少量抗凝血进行细胞培养。加入从红豆中提取的植物血凝素使红细胞集聚并刺激淋巴细胞分裂。这些分裂细胞可被微管抑制剂，如秋水仙素，抑制在有丝分裂。接着，细胞在高渗缓冲液中溶解。中期染色体标本被染色后在显微镜下检查数目上或结构上的改变(图 8-1)。在



图 8-1 经人外周血淋巴细胞培养，Giemsa 显带方法染色的一个中期染色体涂片。在右方显示了一个完整的、非有丝分裂的淋巴细胞核(由 the University of Michigan Clinical Cytogenetics Service 提供)。

细胞周期的 S 期即合成期用氚化的胸腺嘧啶放射性标记 DNA 技术使研究者能够检查染色体复制的时序。20 世纪 70 年代，采用新的染色和显带技术可显著提高染色体形态学的分辨率。

(一) 染色体显带和形态学

目前应用最广泛的 Giemsa 显带技术，即 G 显带，包括对染色体标本轻度的胰蛋白酶处理及随后的 Giemsa 染色。结果产生一种明暗相间的带型，这种带型对每条染色体来说都是独特的，可以明确鉴定每条染色体(图 8-2)。使用这种技术，在每个单倍体都可以观察到大约 350 ~ 550 条带，每条带大约代表 5×10^6 ~ 10×10^6 bp 的 DNA。因为一个基因的长度从 10^3bp(α 珠蛋白)到大于 2×10^6bp(如抗肌萎缩蛋白基因)不等，同时据估计，平均每30 000个 bp 为一个基因，因此，每条带可能代表几个或几百个基因。新技术使我们能够观察到前中期的染色体，在这一时期，染色体比中期更为伸展。这样，观察的分辨率更高(图 8-3)。通过这种技术，每个单倍体可以观察到 850 多条带。

深染的 G 带含有 AT 丰富的 DNA，被认为含有较少的活性基因。这种 DNA 在间期核中呈现出一种浓缩的状态，富含长分散 DNA 序列(long interspersed DNA sequences, LINES)。LINES 是 DNA 的重复区域，不编码表达基因。浅染带，即 R 带(反带)含有 GC 丰富的 DNA，含有许多转录基因。这种 DNA 在间期核中呈现出较为伸展的状态。除了

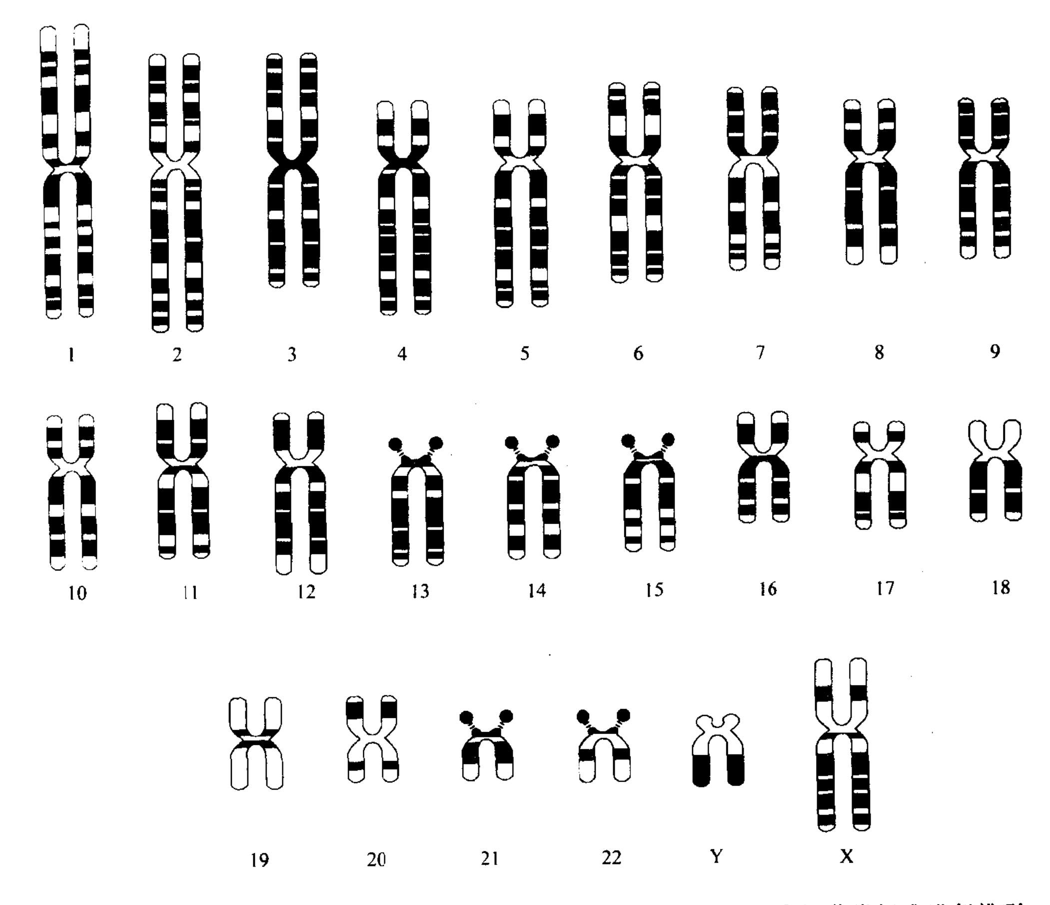


图 8-2 人类染色体模式核型显示 Giemsa 显带模式,按照 1971 年巴黎会议分类标准进行排列和编号(引自 deGrouchy J. Turleau C. Clinical Atlas of Human Chromosomes. 2nd ed. New York : John Wiley & Sons, 1984.)。

转录基因之外,它还含有短分散 DNA 序列(short interspersed DNA sequences, $SINE_S$),包括 *Alu* 序列。染色体上大多数的断裂点和重排被认为是发生在 R 带。

染色体的形态学是由着丝粒(中央狭窄处)的位置来确定的。着丝粒将染色体分为一条短臂(以 p 为标志)和一条长臂(以 q 为标志)(图 8-4)。如果着丝粒位于染色体的中央,这种染色体就是中央着丝粒染色体。如果着丝粒稍微偏离染色体的中央,这种染色体就是亚中着丝粒染色体。如果着丝粒靠近染色体的末端,这种染色体就是近端着丝粒染色体。人类近端着丝粒常染色体(13～15 号、21 号和 22 号染色体)的短臂主要由一个随体和一条细柄组成。这些区域可以形成静止期细胞的核仁,含有核糖体 RNA 基因的大量串联重复拷贝。

有两个结构特征值得特别注意。**着丝粒**(centromere)连接两条**姐妹染色单体**(sister chromatids),保证两者在有丝分裂和减数分裂时进行正确的分离。同时,它也是动粒(kinetochore)形成的位点,并与纺锤体的微管相连,在细胞分裂时调节染色体移动。着丝粒还含有"卫星"DNA,它是一种短串联重复 DNA 序列,通常具有染色体特异性。**端粒**(telomere)位于染色体的末端,它是一种特殊的蛋白质－DNA 结构,含有 TTAGGG 六核苷酸

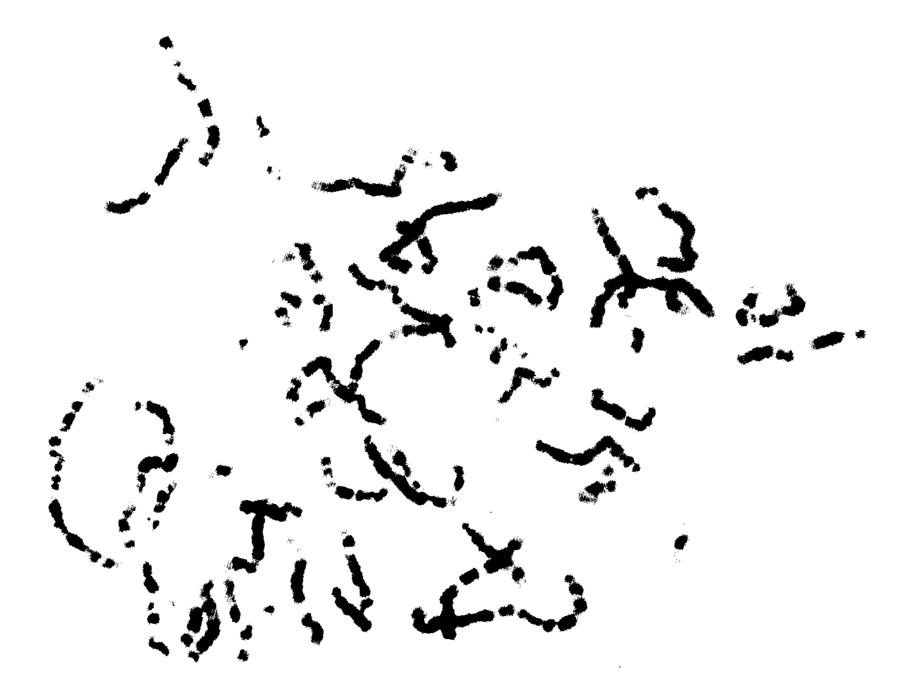


图 8-3 停止在前中期的、外周血淋巴细胞培养的人染色体经 Giemsa 显带技术染色(由 the Unversity of Michigan Clinical Cytogenetics Service 提供。)。

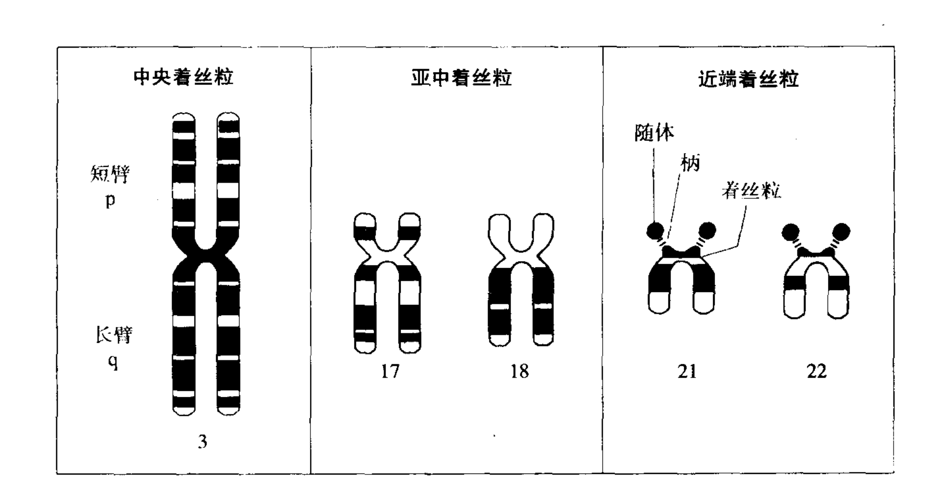


图 8-4 按照染色体大小变化和着丝粒位置分类的人类染色体形态。

重复的延伸序列。端粒可以保护染色体末端不被降解,并防止与其他染色体间的末端融合。端粒长度的缩短与体细胞的老化有关。相比之下,生殖系和肿瘤细胞含有端粒酶,可以使端粒恢复原长。

如图 8-5 所示,每条染色体上的带型按照规则从着丝粒到染色体长、短臂的末端来排号。这样,染色体就可以根据它们的大小,与着丝粒位置有关的形态和带型而得以确定。

在核型中,染色体按照这些标准进行排列(图 8-6)。比较从中期染色体涂片得到的

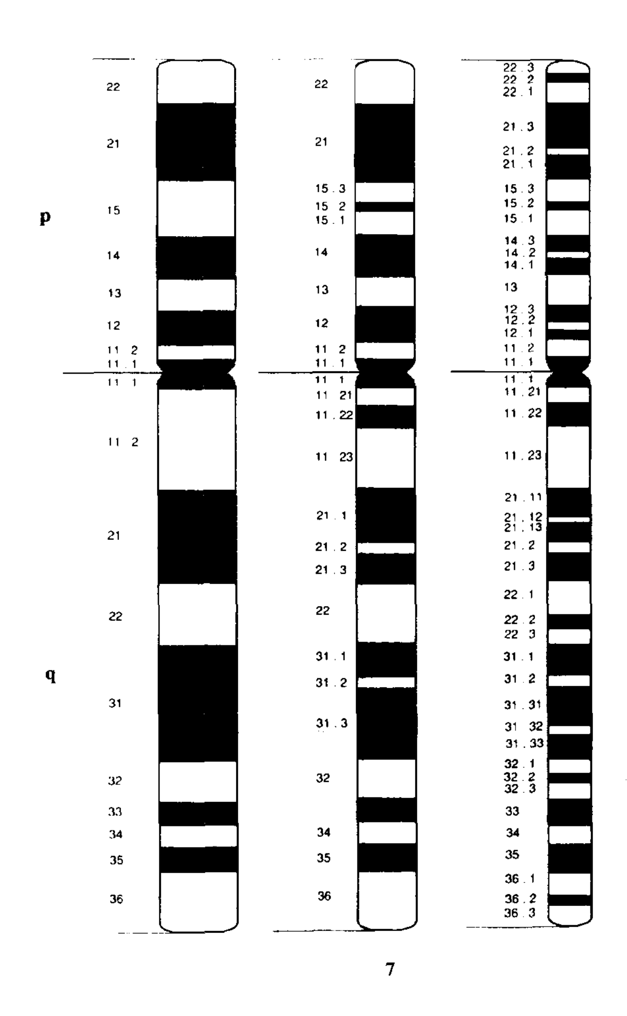


图 8-5 7号染色体的高分辨 G 显带。分辨率为每单倍体组 450 条带(左),550 条带(中)和 850 条带(右)(引自 An International System for Human Cytogenetic Nomenclature. Cytogenet Cell Genet 1985;34:53. ©WS Karger. Basel.)。

人类染色体图像(如图 8-1 和图 8-6)和人类染色体模式核型(如图 8-2 和图 8-4 所示)时,发现两者之间似乎存在差异。在前者,染色体通常像单根的长形腊肠,而在后者,染色体则被描绘成 X 形或 H 形。事实上,经过 S 期的 DNA 复制后,每个染色体被复制,由两条一致的姐妹染色单体组成。在有丝分裂时,姐妹染色单体分离,每一条进入一个子代细胞。按照惯例,人类染色体模式核型显示两条姐妹染色单体虽在着丝粒相连,但两者之间轻度分离。相比之下,在大多数中期染色体涂片上,姐妹染色单体是紧密并置在一起的,因此看起来仿佛是单一结构。

描述核型时,首先列出染色体总数,然后是性染色体组成,接着列出任何染色体数目或染色体形态的异常。表 8-1 列出了目前描述染色体的专有术语。如 47,XY,+21 表示一名 XY 男性有 47 条染色体,包括一条额外的、会引起 Down 综合征的 21 号染色体的拷贝。45,XY,der(14;21) (q10;q10)表示一名 XY 男性有 45 条染色体,其正常的 14 号和 21

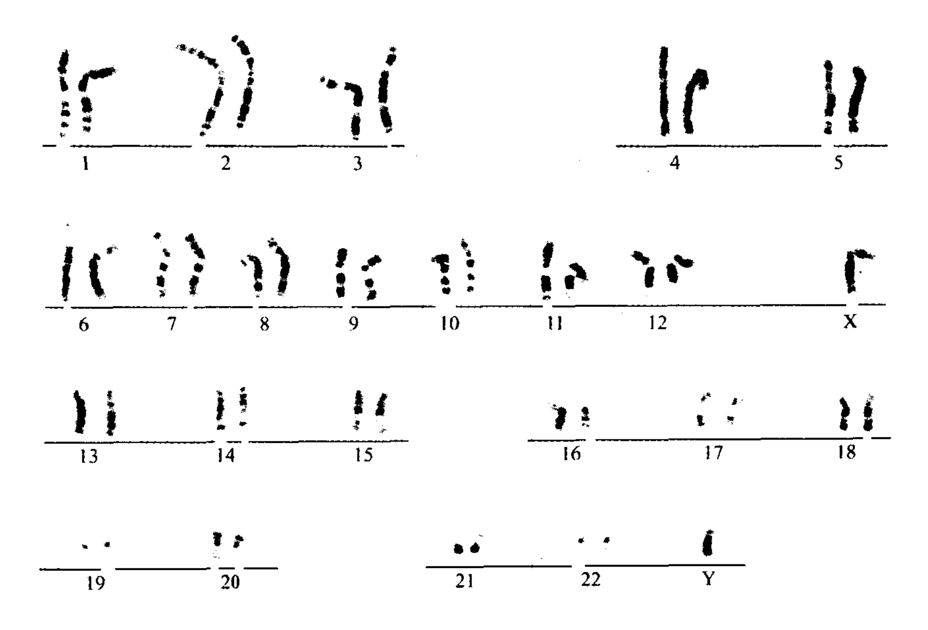


图 8-6 人类 G 显带染色体的核型。染色体按照大小和着丝粒位置排列成组(由 the University of Michigan Clinical Cytogenetics Service 提供。)。

号染色体被 21 号染色体长臂与 14 号染色体长臂易位产生的衍生染色体替代。这就是罗伯逊易位,即着丝粒融合,我们会在后面对它进行讨论。

表 8-1 染色体术语*

A-G	染色体组
1~22	染色体编号
X,Y	性染色体
/	斜线表示嵌合体,如 46/47 表示 46 条染色体和 47 条染色体细胞系的嵌合体
p	染色体短臂
q	染色体长臂
del	缺失
der	衍生,结构重排的染色体
dup	重复
i	等臂染色体
ins	插入
inv	倒位
r	环状染色体
t	易位
ter	末端
+ 或 -	放在染色体号前,这些符号表示多了或丢失一整条染色体,如 +21 表示一条多余的 21 号染色体,如在 Down 综合征。放在染色体号后面,这些符号表示多出或缺失染色体的一部分,如 5p- 表示 5 号染色体短臂部分缺失。然而,del(5p)是更好的术语。

* 引自 Mitelman F, ed. An international system form human cytogenetic nomenclature ISCN(1995). Basel: S. Karger, 1995.

(二) 荧光原位杂交

近年来,细胞遗传学最重要的进展就是**荧光原位杂交**(fluorescence in situ hybridization, FISH)技术的发展。它在鉴别和确定与出生缺陷和肿瘤有关的染色体异常方面得到应用,而且如在第九章讨论的那样,它也是基因定位的有用工具。FISH 技术包括化学修饰的核苷酸标记的特异性 DNA 探针的制备,这些核苷酸可以直接产生荧光,或通过结合荧光标记的报告分子而被检测出来。这些单链的探针不仅像在标准细胞遗传学技术中那样与中期染色体进行杂交,而且还能与非分裂细胞固定的间期核直接进行杂交。标准细胞学需要分裂细胞和阻断在分裂中期的染色体,而 FISH 技术将染色体分析扩展到间期细胞。主要的技术进步还包括具备制备多种探针的能力和数码成像显微镜的改进。

在 FISH 技术中,有几种不同类型的探针可供使用(图 8-7 和图 8-8)。在噬菌体或 YAC 内克隆特异性 DNA 片段和通过 PCR 进行 DNA 扩增可以制备得到基因座特异性探针。作为基因制图方面如人类基因组计划(第十章)努力的结果,可以得到越来越多这样的探针。着丝粒重复探针利用了人类染色体着丝粒处短的、串联重复的染色体特异性 DNA 序列。这些重复探针可在中期染色体和间期核中都产生强烈的信号,因此它们在鉴定特异性染色体方面非常有用。许多此类探针现已商品化。全染色体探针通过对来自特异性染色体流式分检库或仅含一条人类染色体的杂种细胞的 DNA 进行 PCR 扩增而制备。这个 DNA 序列的混合物与一条染色体上的多个位点同源,因此,在中期和间期核

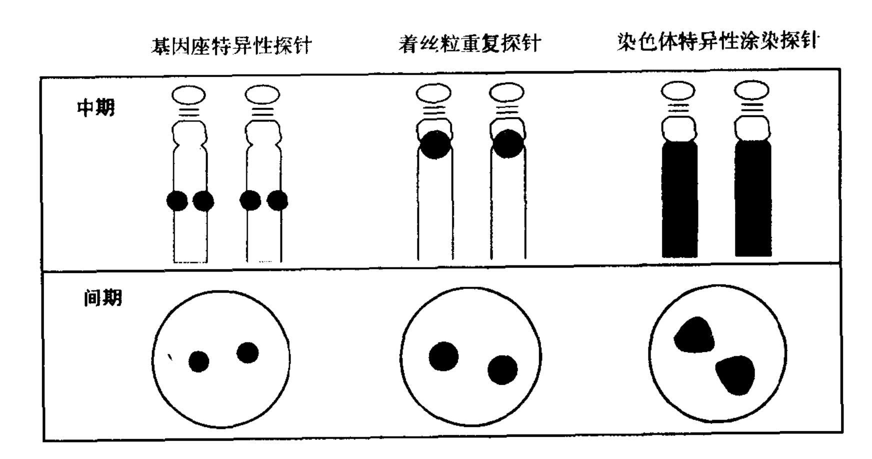


图 8-7 使用三种不同类型探针的 FISH 分析图示。上方图板显示来自中期涂片的一条常染色体的两个拷贝与三种不同类型探针杂交的假想表现。基因座特异性探针在相关位置给出一个尖锐的分离信号。注意每一条姐妹染色单体的杂交都能被看见。Alphoid,也叫 α 卫星,或着丝粒重复探针在着丝粒附近给出一个更大的、更扩散的,但非常强烈的信号。而涂染探针将整个染色体一致着色。相关染色体的拷贝数能通过这些探针中任何一个检测出来。下方图板显示与每一种探针杂交后中期核的表现。在这一环境中,基因座特异性探针和着丝粒探针对查出的信号数目给出最佳的分辨率(引自 Emanuel BS. The use of fluorescence *in situ* hybridization to identify human chromosomal anomalies. Growth Genet Horm 1993;9:6~12.)。

内,整条染色体得以被涂染。通过使用不同的荧光素,在一个中期染色体涂片上鉴别几条不同的染色体成为可能。以不同的比例混合 5 种基本荧光素和应用复杂的成像技术,在一个涂片上 24 种不同的染色体都有可能得到染色(图 8-8 见后彩页)。进一步的发展是生产区域特异性探针的染色体区域的微切割。对单个中期染色体某一特异性区域进行微切割,继之进行 DNA 的 PCR 扩增,使我们无需进行费时的克隆过程就能制备这种区域特异性探针。

FISH 技术上的其他改进包括高分辨"纤维 - FISH"和比较基因组杂交。前者在基因制图上很有用,将在第九章进行讨论,后者涉及到与两种不同颜色标记的等量受检(test) DNA 和参考(reference)DNA 的正常中期涂片进行联合杂交。这项技术在肿瘤细胞中检测和定位缺失或扩增方面特别有用。

FISH 技术在诊断上的应用包括发现和确定在中期涂片上染色体数目和结构的异常。染色体数目异常在任何组织来源的间期细胞内都能很容易地被检测出来。因此,这些技术在先天性染色体疾病、某些白血病、淋巴瘤、实体瘤的临床评估和产前诊断上有着重要的应用。特殊病例将在本章以后的部分进行讨论。FISH 还可以应用于单倍体细胞如精子,从而获得有关配子中染色体异常频率的信息。

二、染色体异常

染色体异常可能是数目的,也可能是结构的。在讨论染色体数目异常时,明确"倍体"(ploidy)是很重要的。**整倍体**(euploid)指染色体数是正常单倍体配子中所含染色体数 N 的整倍数。一个正常体细胞内的染色体数是**二倍体**(diploid),即 2N。整倍体染色体组成不一定都是正常的,如三倍体(69 条染色体)就是一种常发生在自发流产中的整倍体染色体数目异常(图 8-9)。**非整倍体**(aneuploid)指任何不是整倍体的染色体数目。对全身的染色体疾病来说,它通常指某一条染色体多了一个额外的拷贝,即**三体性**(trisomy)(如

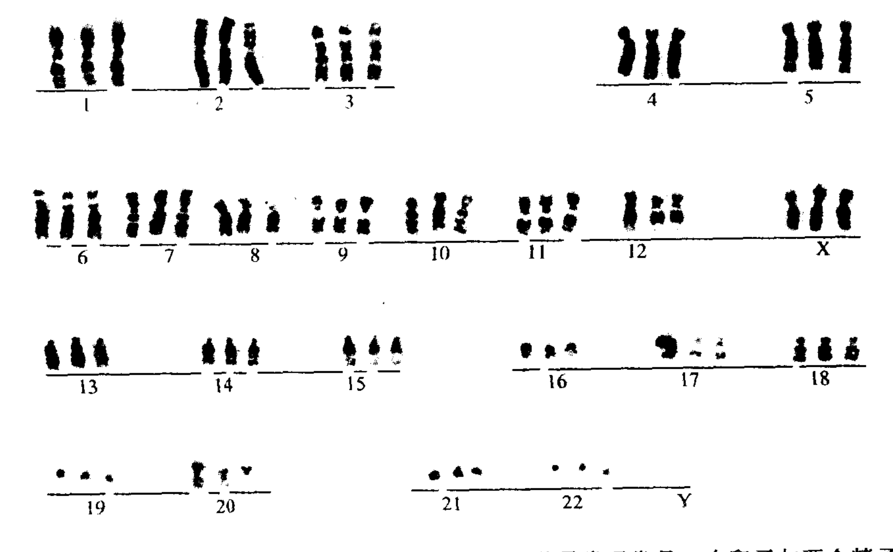


图 8-9　三倍体新生儿的外周血淋巴细胞核型。这种异常通常是一个卵子与两个精子受精的结果(由 the University of Michigan Clinical Cytogenetis Service 提供)。

21 三体性 Down 综合征);或者缺失了某一条染色体,即**单体性**(monosomy)(如 45,X Turner 综合征)。

(一) 数目异常

非整倍体是**不分离**(nondisjunction)的结果,即在细胞分裂时染色体不能正常分开。它可以发生在减数分裂或有丝分裂时。虽然引起不分离的原因仍不知晓,但人们认为减数分裂不分离的风险随母亲生育年龄的增高而上升。当不分离发生在受精之后,即有丝分裂不分离,将会见到嵌合体,即存在两种或两种以上的不同染色体的细胞群。

如图 8-10 所示,减数分裂不分离可发生于第一次减数分裂或第二次减数分裂,从而带来十分不同的后果。如果不分离发生在第一次减数分裂时,形成的配子将会包含双亲两个未能分离的染色体(父源和母源)或一个也没有。当不分离发生在第二次减数分裂时,配子将会含有双亲之一染色体的两个拷贝(父源或母源)或一个也没有。因此,在第一次减数分裂时发生的减数分裂不分离所产生的配子受精后,其后代将同时含有一对来自于祖父和祖母的同源染色体拷贝。相比之下,当不分离发生在第二次减数分裂时,后代将含有一个来自于祖父或祖母的染色体的两个拷贝。注意,由于在减数分裂的四-染色单体阶段发生重组(第二章),所以这两个拷贝在遗传上不是一致的。还要注意,第一次或第二次减数分裂发生的不分离也会产生单体性合子。

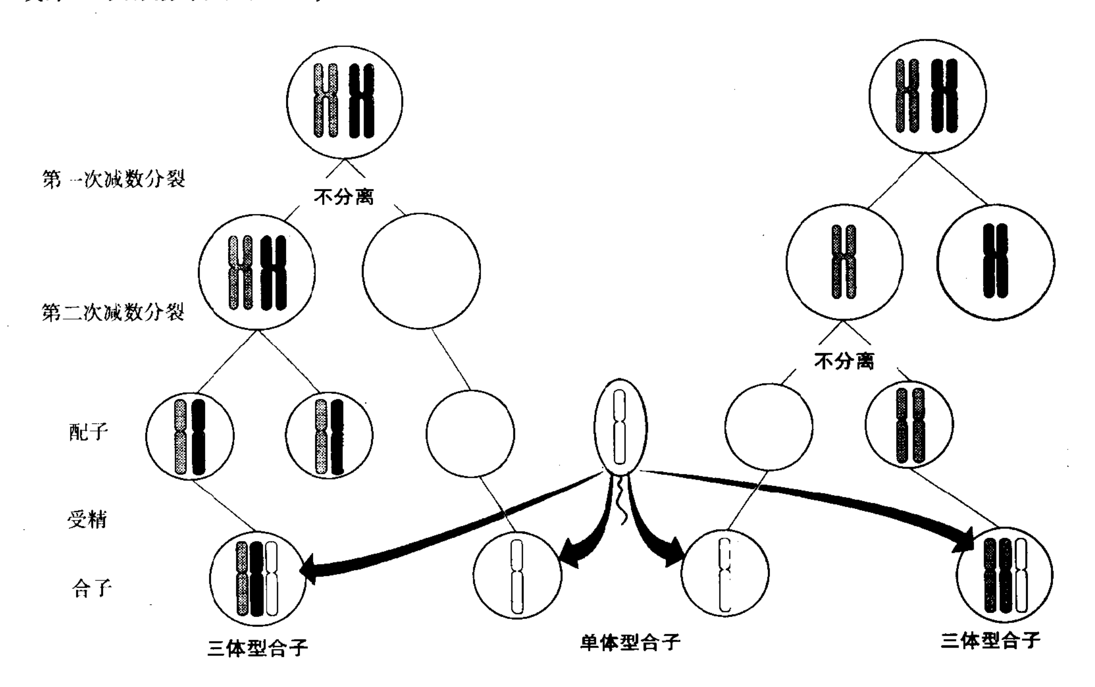


图 8-10 发生在第一次或第二次减数分裂的染色体不分离的后果。详情见书。

(二) 结构异常

染色体的结构异常可能涉及一条或两条(或更多)的染色体(图 8-11)。当仅涉及一

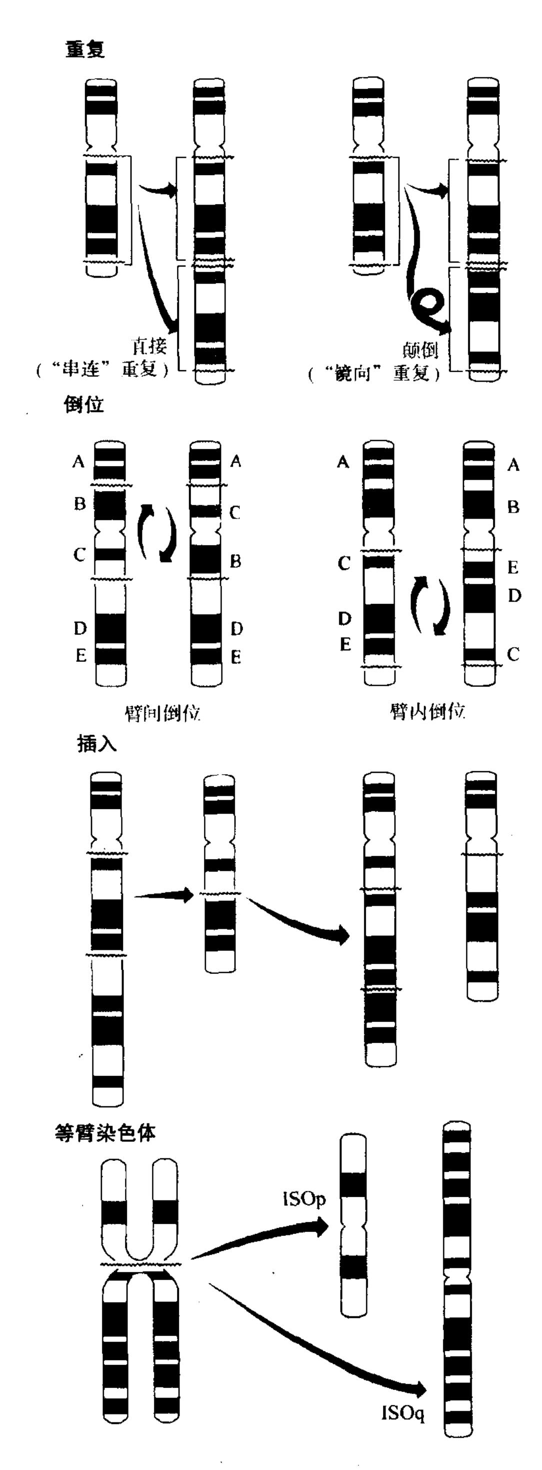


图 8-11　染色体结构异常。倒位涉及单个染色体的两次断裂。如果断裂包括了着丝粒，为臂间倒位；如果没有，为臂内倒位。插入至少由两条染色体的三次断裂形成。核型表现出不平衡的后代具有插入区域的重复或缺失。等臂染色体是由着丝粒异常分裂造成的，可以引起短臂重复和长臂缺失（ISOp）或者长臂重复和短臂缺失（ISOq）。后者不在 X 染色体发生。

条染色体时,该染色体可能有部分缺失或重复、一次臂间(涉及着丝粒)或臂内(不涉及着丝粒)倒位、环状染色体或等臂染色体形成。涉及两条或多条染色体的结构异常可能会导致一条染色体的物质插入另一条染色体,或更常见的两条或多条染色体间染色体物质的易位或交换。染色体易位两种主要的形式已确定:相互易位(图 8-12A)和罗伯逊易位,

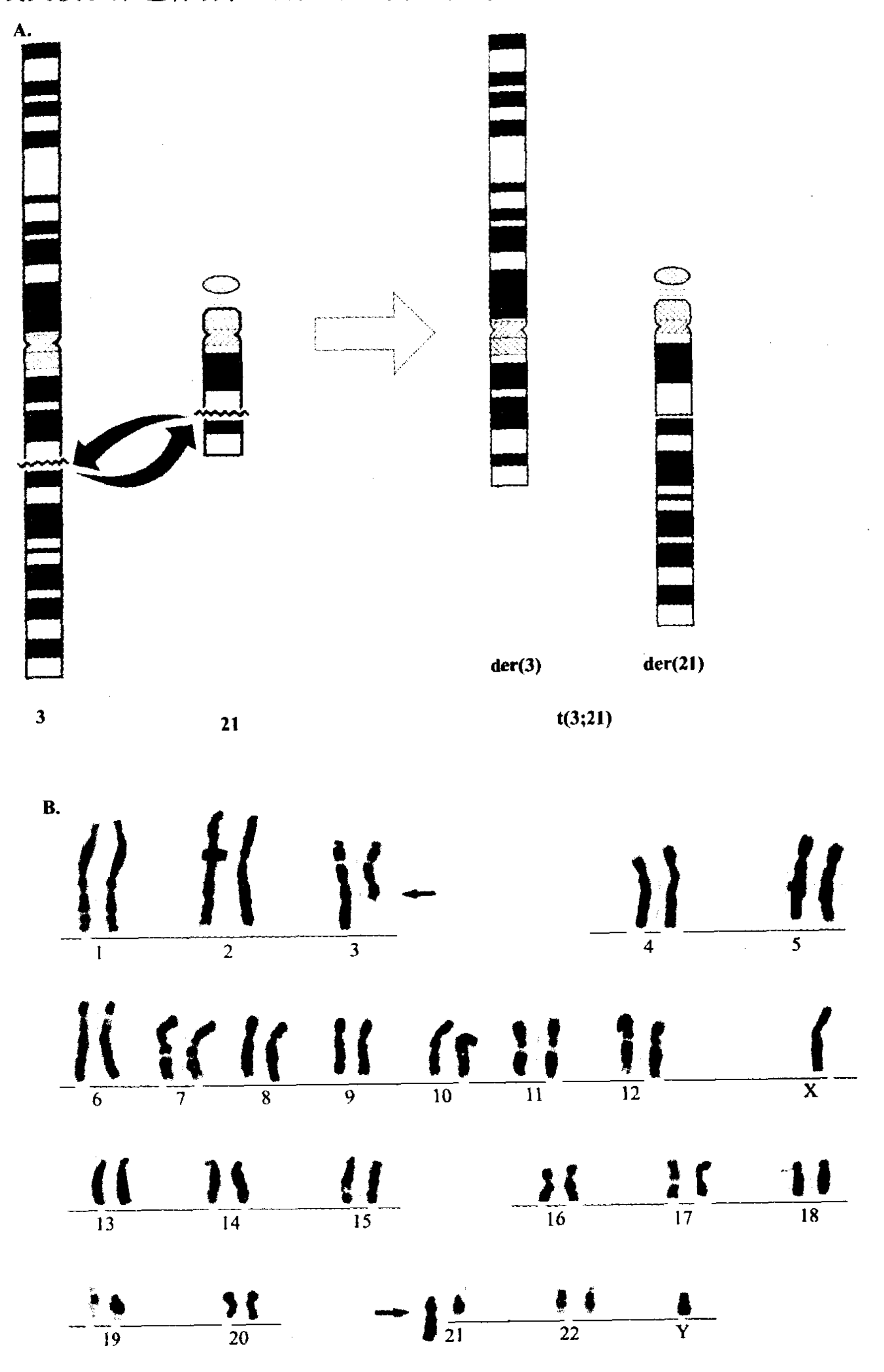


图 8-12 3 号染色体和 21 号染色体之间的相互易位。在图中,每条染色体长臂上的断裂点用波浪线表示(A)。在核型中,易位染色体[der(3)和 der(21)]用箭头指示(B)。

即着丝粒融合(图 8-13A)。在这两种易位中,如果断裂和重接没有造成重要的染色体物质的丢失和基因的损伤,那么携带这种易位的个体在临床上是正常的。但是,该个体产生染色体不平衡的后代的风险增高。这在图 8-14 和 8-15 中加以解释。

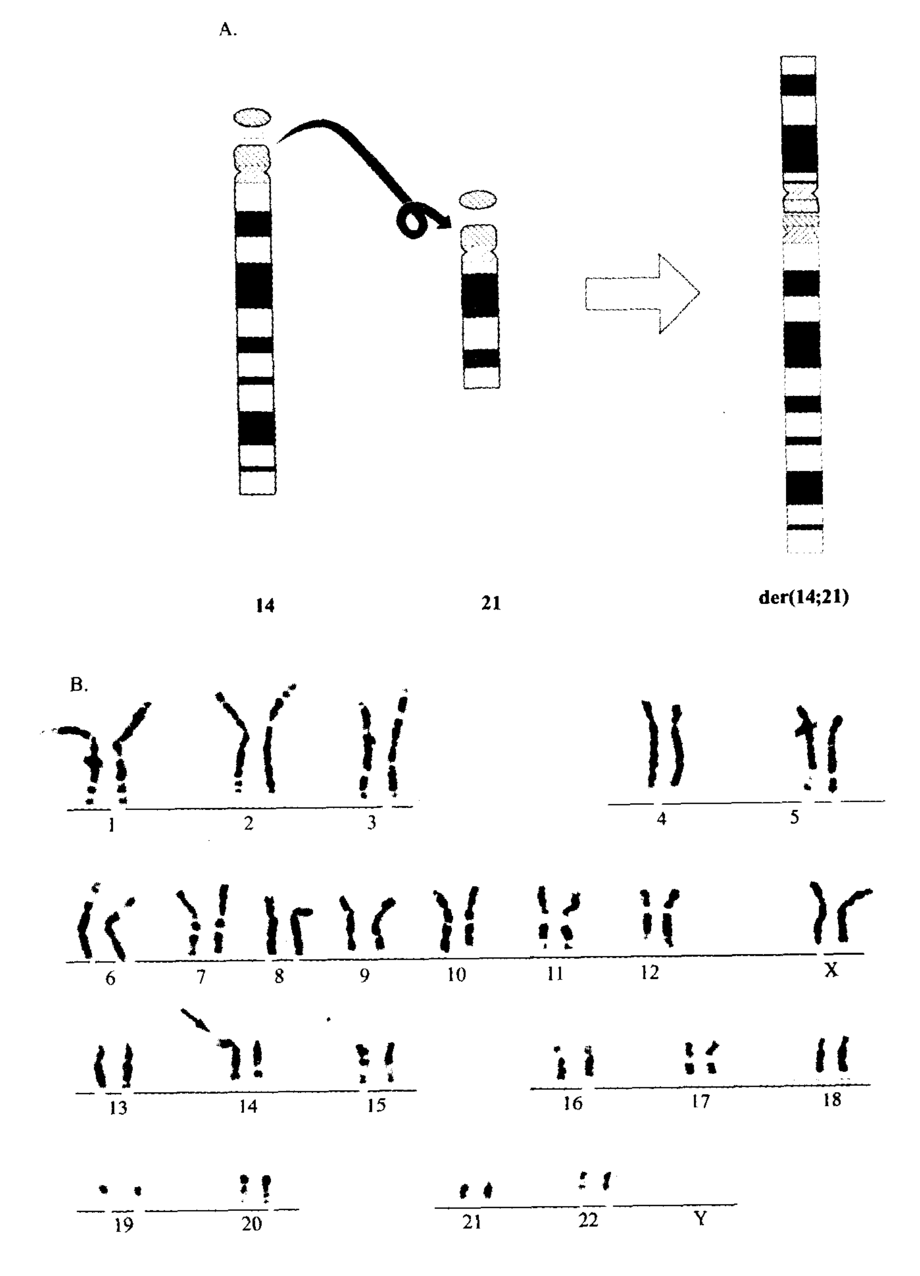


图 8-13 涉及 14 号和 21 号染色体长臂的罗伯逊易位,即着丝粒融合。在核型中,易位染色体用箭头指示。

在图 8-12 和图 8-14 所示的例子中,3 号染色体和 21 号染色体之间存在涉及两者部分长臂的相互易位。在减数分裂时,它们会形成一种交叉形或四射体样结构(图 8-14),同源染色体能够以几种不同的方式进行分离。在对位分离(alternate segregation)中,两条

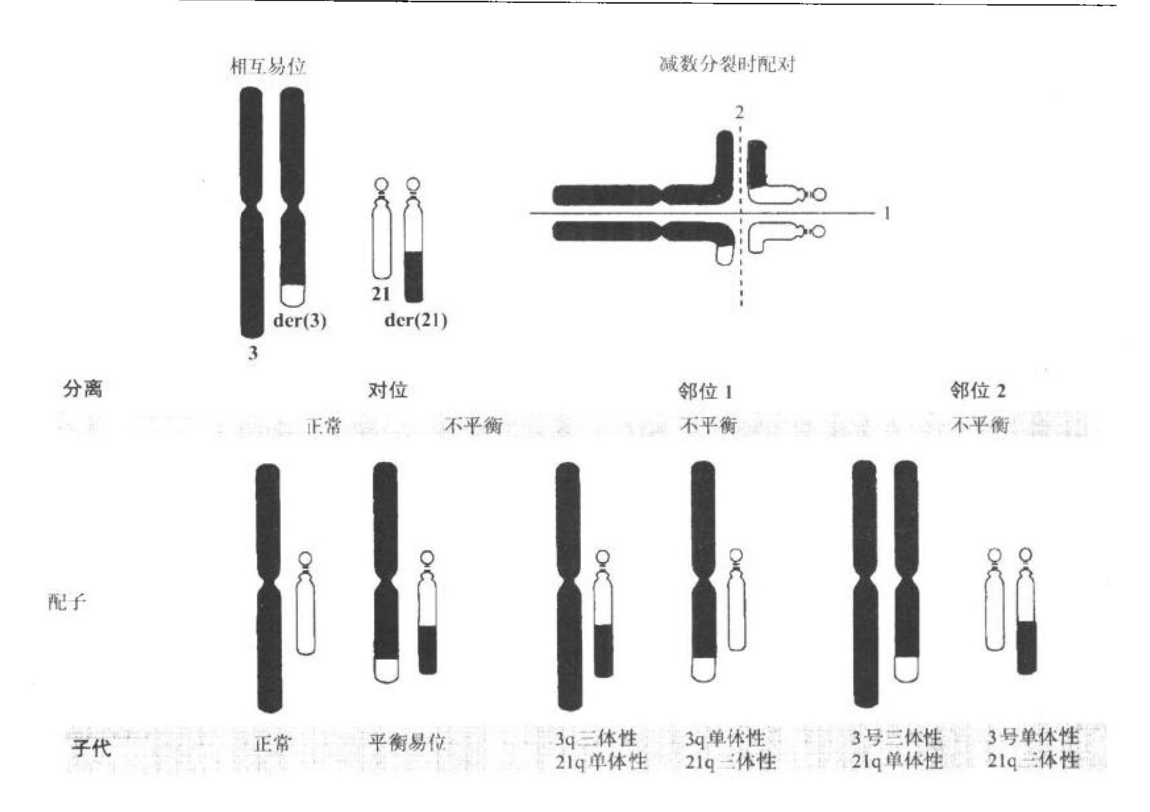


图 8-14 涉及 3 号和 21 号染色体的相互易位中减数分裂的分离。详情见书。

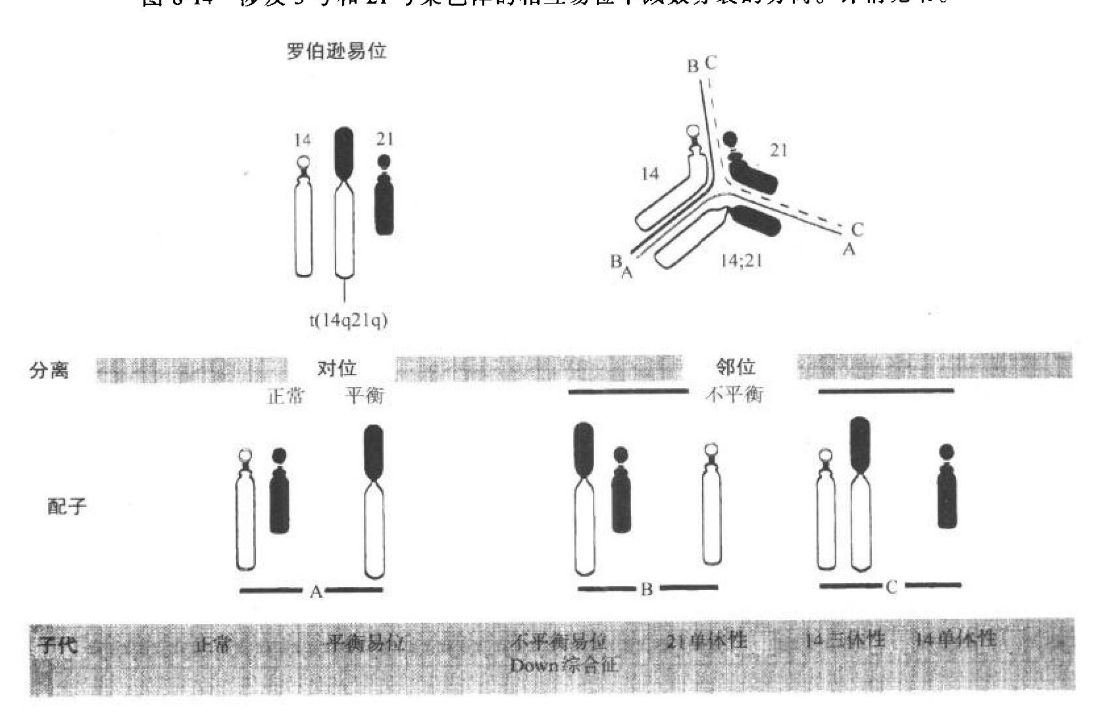


图 8-15 涉及 14 号和 21 号染色体的罗伯逊易位中减数分裂的分离。详情见书。

正常的染色体进入一个子细胞,而易位的染色体进入另一个子细胞。在这种情况下形成两种配子:正常的配子和携带平衡相互易位的配子。正常的配子受精产生正常的后代,而平衡易位的配子受精产生和双亲一样的、临床上正常的、平衡易位后代。然而,如果减数分裂时染色体的分离导致产生含有一条正常的染色体和一条易位染色体的配子,如在称为邻位(adjacent)1 和邻位 2 的分离中,那么,不平衡的染色体组成通过配子而传递。

在邻位 1 分离中,同源的着丝粒彼此分离,不平衡的配子受精后,产生的后代或者含有 3 号染色体长臂末端的一个额外拷贝并缺失了 21 号染色体长臂的一部分,或者缺失 3 号染色体长臂的一部分并含有来自 21 号染色体长臂的额外的物质。邻位 2 分离中,同源的着丝粒不分离,产生的配子受精后导致两条染色体中一条染色体几乎完全为单体性,另一条染色体为部分三体性。即使在胚胎早期,这种染色体异常程度与存活也是不相容的,导致自发流产,所以罕见这样的异常。

很明显,如果某种分离方式(如邻位 2)产生染色体不平衡的程度比其他方式更严重,那么该方式在存活的后代和胎儿中就比较少见,因为他们存活的可能性较小。然而,在减数分裂中,不同的分离方式发生的频率似乎也不同。应用仓鼠卵细胞同人类精子进行体外受精的技术,我们可检查出男性精子中杂合性平衡易位的核型。大约 50% 的精子的核型是平衡的,它们是对位分离的结果。正常和平衡易位的精子的频率相同。大部分的染色体不平衡的配子是由邻位 1 分离造成的。不平衡的核型在精子中比在胎儿和活产儿中更多见,这再次说明存在对染色体异常的合子的宫内选择。

图 8-13 显示 14 号染色体和 21 号染色体之间典型的罗伯逊易位,即着丝粒融合。这样的易位发生在 D 组和 G 组染色体间,即 13、14、15、21、22 号染色体。它们都是近端着丝粒染色体,其短臂由一个随体和细柄组成,含有核仁组织区(nucleolar organizing region, NOR)和串联重复的核糖体 RNA 基因。在减数分裂中,这些染色体趋于簇集,比其他染色体更常发生断裂和重组,造成这些染色体的短臂丢失和长臂之间互相融合。在分子水平,这些融合是不平衡的,因为它们缺失了两条染色体的短臂;但在表型水平,他们是平衡的,因为核糖体 RNA 基因丰富的储备使这些短臂丢失不会产生严重的后果。这些短臂丢失在临床上也是不明显的,因为携带这种平衡罗伯逊易位的个体在临床上是正常的。然而,他们具有产生临床表现异常的不平衡易位后代的高风险。图 8-15 所示为含有 45 条染色体以及 14 号染色体和 21 号染色体长臂融合的个体。在减数分裂中,涉及的染色体会排列成三射体的构象。如图 8-15 所示,这些染色体在减数分裂中有三种可能的分离方式。

在减数分裂的对位分离中(如 A 线所示),一个配子接受一条正常的 21 号染色体和一条正常的 14 号染色体,而另一个配子接受那条衍生的、即携带 14 号染色体长臂和 21 号染色体长臂的易位染色体。这两种配子与正常配子受精后,会分别产生一个正常的 46 号染色体的核型和一个 45 条染色体的平衡易位的核型,后者与易位携带父母的核型相同。这些后代在临床上是正常的。

另一方面,如果发生如 B 线所示的邻位分离,则一个配子将接受一条正常的 14 号染色体,但缺失一条 21 号染色体。和正常配子受精后将产生 45 条染色体的核型,仅含有

21 号染色体的一个拷贝，因此是 21 单体性。它通常是不能存活的。形成的另一个配子会含有 21 号染色体的正常拷贝和易位染色体。与正常配子受精后，该合子含有 46 条染色体，包括 21 号染色体两条正常的拷贝，一条正常的 14 号染色体拷贝以及包含 14 号和 21 号染色体长臂的易位染色体。在功能上，该个体有 21 号染色体长臂的三个拷贝，因此具有 Down 综合征的临床表型。他（假定为男性）的核型是 46，XY，+21，der，(14;21)(q10;q10)。

最后是 C 线所示的邻位分离。一个配子接受 14 号染色体的正常拷贝和易位染色体。这样的配子与正常配子受精后产生含有 46 条染色体的合子和功能性 14 号染色体三体性，这样的个体不能存活。另一个配子将接受正常的 21 号染色体但没有 14 号染色体。与正常配子受精后，合子仅含有 45 条染色体，缺失了 14 号染色体的一个拷贝。14 单体性也是不能存活的。

因为 21 单体性、14 单体性和 14 三体性都是不能存活的，所以在一个 14q 和 21q 易位的双亲的后代中会出现三种可能的结果：正常的、平衡易位和引起 Down 综合征的不平衡易位。然而，这三种类型发生的频率并不相同。相反，当母亲是 14q;21q 易位的携带者时，大约 10% ~ 15% 的后代患有 Down 综合征，因为他们携带正常 21 号染色体的两个拷贝和一条易位染色体。大约一半的后代是和母亲一样的易位携带者，另外 40% 后代的染色体是正常的。出于某种未知的原因，当父亲是易位携带者时，孩子的发病风险相对较低，低于 5%。存活的后代中 21 三体性的频率低于预期水平，这反映了三体性胎儿高频率的自发流产。

三、染色体异常的临床后果

（一）自发流产

从数量上讲，染色体异常的主要临床后果发生在出生前，大约 50% 或更多的自发流产胎儿有严重的染色体异常。因为大约 15% 的已知妊娠以自发性流产而告终（通常发生在妊娠头三个月以内），所以这就说明，至少有 7.5% 的妊娠有严重的染色体异常。而且，现在据估计，50% 的妊娠以自发性流产而告终（其中，许多流产发生的时间很早，只有用激素测试才能检查出来），这表明至少有 25% 的妊娠发生了严重的染色体异常。

在流产胎儿中发现的主要染色体异常类型如表 8-2 所示。就类型来说，三体性是最常见的异常。对个体而言，最常见的异常是 X 单体性（45，X）、三倍体和 16 号染色体三体性。有趣的是，活产儿中未见有 16 三体性，三倍体也很罕见，而且它们易在婴儿期早期死亡；而 45，X 可见于活产女婴中，但仅有约为 1/5 000 的发病率，它是引起 Turner 综合征的原因之一。因此，携带严重染色体异常的妊娠失败率高。

大约 9% 的染色体异常的流产胎儿和大约 2.5% 的死产儿表现有 13、18 和 21 三体性，但活产儿仅有 1% 表现为 13、18 和 21 三体性，这又一次说明了染色体异常的合子的高死亡率。染色体异常在自发流产胎儿中的频率（50%）要高于死产胎儿（定义为孕 28 周后死亡）（5%），这表明大多数死亡发生在妊娠早期。

表 8-2 在染色体异常的流产胎儿中不同异常的相对频率*

类型	发生率	
三体性		52
16	15	
13,18,21	9	
XXX,XXY,XYY	1	
其他	27	
45,X		18
三倍体		17
四倍体		6
其他		7
总和		100

* 引自 Hassold TJ: Chromosome abnormalities in human reproductive wastage. Trend Genet 2:105 ~ 110,1985

(二) 出生缺陷

先天性染色体异常的第二大临床后果是出生缺陷。在新生儿调查中发现的各种染色体畸变的频率列于表 8-3 中。平衡易位是最常见的,发病率为 1/500。它通常不会引起临床异常,但能使成年携带者容易生出染色体不平衡的后代。然而,其他各种不平衡的染色体构成也并不少见,而且染色体异常产生的负担是很明显的。0.5%以上的新生儿带有严重的染色体异常。虽然这些染色体异常的临床表现不尽相同,但普遍的特点是**生长发育迟缓**(growth retardation)、**智力低下**(mental retardation)和**特征性异常体征**(specific somatic abnormalities)。

表 8-3 新生儿中染色体异常的发生率*

	数目	大致发生率
28 580 名男性的性染色体异常		
XXY	30	1/1 000
XYY	26	1/1 000
其他	17	1/1 700
总和	73	1/400
14 976 名女性的性染色体异常		
45,X	2	1/7 500

续表

	数目	大致发生率
XXX	13	1/1 200
其他	5	1/3 000
总和	20	1/750
43 556 名婴儿的常染色体异常		
+D(13 三体性)	3	1/15 000
+E(18 三体性)	4	1/11 000
+G(21 三体性)	45	1/900
其他三体性	2	1/22 000
总和	54	1/800
结构重排		
平衡	81	1/500
非平衡	21	1/2 100
总和	102	1/400
染色体异常总和	249	1/170

* 表中数据来自苏格兰、丹麦、加拿大和美国的六项调查。引自 Jacobs PA, Melville M, Ratcliffe S, Keay AJ, AJ, Syme J: A cytogenetic survey of 11,680 newborns infants. Ann Hum Genet 37:359~376,1974. 剑桥大学出版社。

1. Down 综合征

最先得到公认和最常见的严重染色体疾病之一就是 Down 综合征,通常由 21 号染色体的一个多余拷贝,即 21 三体性所致。它的临床特征包括生长发育迟缓,不同程度的智力低下和包括头面部特征在内的一系列异常体征,这使得它在过去常被不幸地称为先天愚型。这些体征包括脸部和枕骨扁平,双眼上斜,内眦赘皮,巨舌和小耳(图 8-16)。



图 8-16 Down 综合征女孩。注意其直发、苍白脸色、巨舌和张着的嘴,双眼上斜,内眦赘皮。

另外，Down 综合征患儿通常肌肉松软或肌张力低下，40%有明显的先天性心脏病（通常是房室通道缺损），5%有严重的胃肠畸形，包括十二指肠狭窄。他们有高于正常人 15 到 20 倍的白血病发病风险，增高的感染易感性，增高的白内障和甲状腺功能低下的发病率以及成熟前衰老的征兆。事实上，所有存活至 35 岁以上的 Down 综合征患者都出现 Alzheimer's 病的病理表现。Down 综合征儿童面临的最严重的健康问题是先天性心脏病。随着针对这些问题的内科和外科治疗的改进，以及护理这些 Down 综合征患者的态度的改变，他们平均生存期已从 1930 年的 9 年上升至今天的差不多 60 年。

95% Down 综合征患者的染色体基础都是继发于减数分裂时染色体不分离所致的 21 三体性。通过 DNA 多态性分析，能够在大部分家庭中判定不分离发生在双亲中的哪一方以及发生在减数分裂的哪一步（图 8-17）。在 95%的病例中，这条多余的 21 号染色体来自于母亲。在这些病例中，80%是发生在减数分裂Ⅰ期。人们很早就认识到母亲的年龄与其生育 21 三体性孩子（其他三体性，如 13 和 18）的风险率之间有很强的关联（图 8-18），但是继发于易位的 Down 综合征则无此关联。然而，这种关联的基础仍属未知。

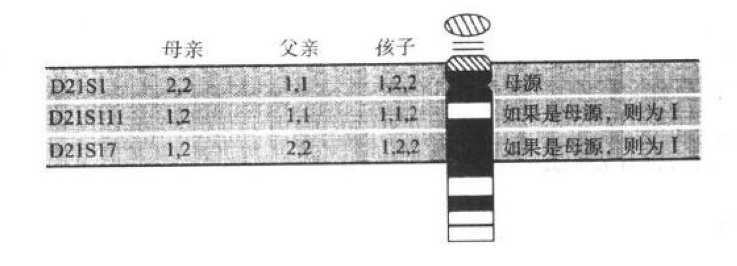


图 8-17 运用 DNA 多态性确定在 Down 综合征中不分离的位置。D21S1、D21S111 和 D21S17 是 21 号染色体上无名的多态性标记位点。D21S1 识别的多态性仅与母亲染色体不分离相关，但无法区分减数分裂Ⅰ或Ⅱ，D21S111 和 D21S17 与父亲的减数分裂Ⅰ或Ⅱ、母亲的减数分裂Ⅰ相容，但与母亲减数分裂Ⅱ不相容。这些结果与母亲减数分裂Ⅰ的染色体不分离最一致。

大约 4%的 Down 综合征病例是非平衡易位的结果。出生时母亲年龄小于 30 岁的 Down 综合征患儿有 9%为不平衡易位，而出生时母亲年龄大于 35 岁的患儿只有不到 2%为不平衡易位。大约 60%的这种易位涉及一条 D 组染色体（13，14 或 15）长臂融合到 21 号染色体长臂，其中大部分是 14 号染色体。大约一半的这种不平衡易位是患儿父母之一（通常为母亲）在配子形成时新发生易位的结果，其余的是遗传下来的易位的结果。后者很重要，因为它提示其他的家庭成员也可能是易位携带者，生出 Down 综合征患儿的风险高。大约 40%的易位仅涉及 G 组染色体，即 21 和 22 号染色体。其中，90%是明显的 21q;21q 的易位（其实是 21q 等臂染色体），这些几乎都是新发生的遗传事件。在父母携带 21q;21q 易位的情况下，惟一能存活的个体只能是 Down 综合征。最后，不到 3%的 Down 综合征患者是嵌合体，他们含有一个 47，+21 的细胞系和一个正常的整倍体细胞系。这些个体的 Down 综合征表现较轻。

对一对已有一个 21 三体性患儿的父母来说，从经验上看，再发风险为 1%。造成这种风险增高的原因尚不清楚。对平衡易位的携带者来说，如果母亲是携带者，孩子的受累风险大约为 10%～15%；若父亲是携带者，则受累风险小于 5%。另外，如果这种易位

是遗传来的，就有必要认识到其他亲属也有类似的发病风险。平衡易位携带者的一级亲属有50%的机率也是该易位的携带者。

21号染色体是最小的染色体之一，含有大约1.7%的细胞总DNA和大约1 500个基因。通过对一些仅是21号染色体部分三体性患者的研究，人们认为只要21号染色体长臂远端1/3，特别是q21.2至q22.3出现一个多余的拷贝就可以引起Down综合征。目前，大量的基因已被定位于21号染色体的这一区域，并且在这些基因中有几个在21三体性中存在基因剂量效应。淀粉样前体蛋白(amyloid precursor protein)、超氧化物歧化酶-1、某些嘌呤合成酶、人类癌基因*ETS-2*和α-A晶状体蛋白(crystallin，一种晶体内蛋白)的基因有着特别重要的意义(图8-19)，因为Down综合征患者在生命早期就出现Alzheimer's病，通常他们体内的嘌呤水平高，具有增高的急性白血病发病风险和较高的白内障发病率。但是，这些表型特征是否与21号染色体长臂上特异性基因的三体性直接相关仍待进一步证实。

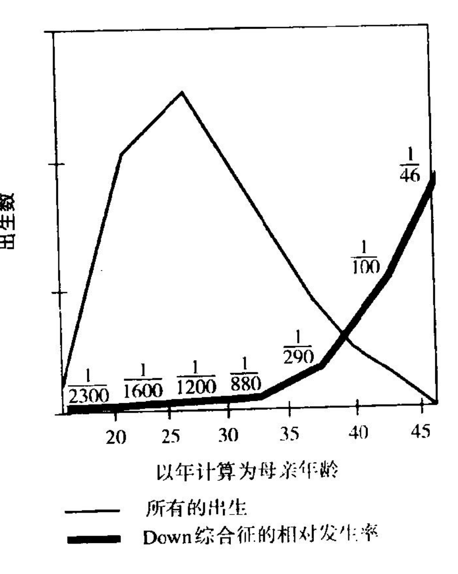


图8-18 母亲年龄和由21三体性引起的Down综合征发病率之间的关联(引自Hamerton JL. Human Cytogenetics Vol. Ⅱ, New York: Academic Press, 1971.)。

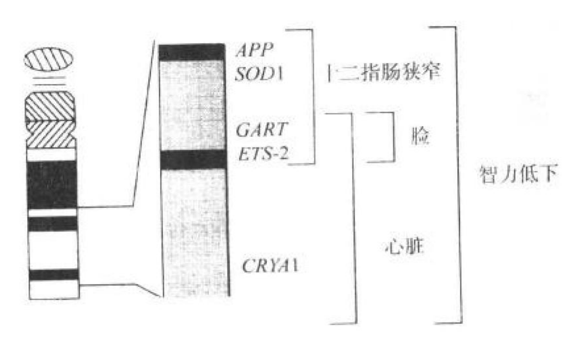


图8-19 以21号染色体部分三体性患者的分子和细胞遗传系分析为基础得到的人类21三体性的表型图。只有少数基因定位于21q21.2～21q23.3(引自Epstein CJ. Down syndrome (trisomy 21). In: Scriver CR, Beaudet. AL. Sly WS, Valle D, eds. The Metabolic and Molecular Bases of Inherited Disease. 7th ed. New York: McGraw-Hill, 1995:749～794.)。

2. 其他常染色体异常

其他严重的染色体异常包括一些三体性和部分三体性。例如13三体性，在活产儿中发病率约为1/15 000。这些患儿中大约一半在生命的第一个月就死亡了。在特征上，他们具有正常脑的发育缺陷，前大脑皮质形成缺如，称为前脑无裂畸形。同时还有相关的头颅和面部畸形，包括严重的唇裂、腭裂和面部中段畸形(图8-20)。特有的凿孔状头皮边缘

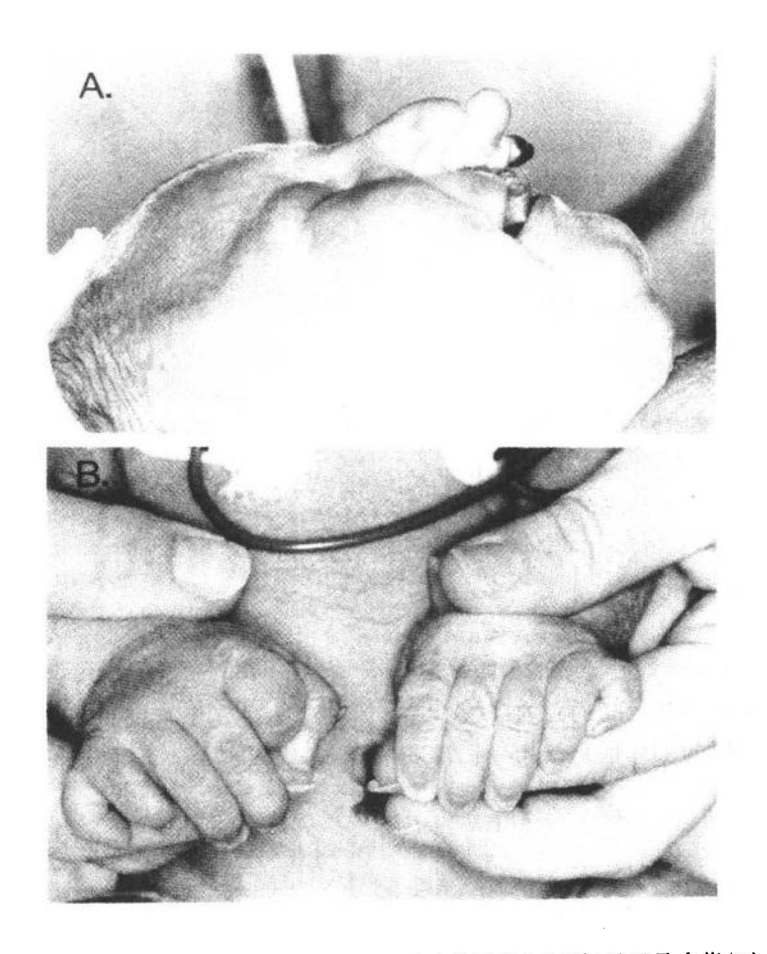


图 8-20 13 三体性的婴儿。注意面部畸形和严重的唇裂和腭裂(A)以及多指(B)。

缺陷和多趾(指)是这一综合征的特点。和其他严重的染色体异常一样,这些患儿生长缓慢,发育严重迟缓。

其他染色体异常涉及结构缺陷,如部分三体性和部分单体性,即某条染色体多余或缺少部分片段,导致多个基因的剂量改变。**微缺失综合征**(microdeletion syndrome)或**相邻基因综合征**(contiguous gene syndrome)通常是散发的,但偶尔也有家族性。它具有一致的但复杂的表型,这一表型与一微小染色体(通过 < 5Mb)缺失有关。例如,21q11.2 中间缺失的半合子与一种可变的 CATCH22(心脏缺陷、异常面容、胸腺发育不全、腭裂、低钙和 21q11 缺失)临床综合征有关。还包括一系列的疾病,如 DiGeorge 综合征[胸腺发育不全,甲状旁腺发育不全、面部畸形和锥体(心室流出道)心脏缺陷]和腭心面综合征(腭裂、鼻部畸形、发育延迟和锥体缺陷)。作为一类染色体异常,这些 22q11 缺失的发病率为 1/5 000,在 5% 的先天性心脏缺陷中发挥作用。位于此区域的基因在第三和第四咽囊包括甲状腺、甲状旁腺和锥体的发育中一定很重要。虽然微缺失有时通过高分辨前中期染色体分析可以检查出来,但 FISH 分析要更敏感而且更实用。在一个受累个体中靶向关

键的染色体区域的探针将在一个中期染色体涂片上仅与两条同源染色体中一条杂交，所以在间期核中仅有一个而非两个杂交信号（图 8-21）。

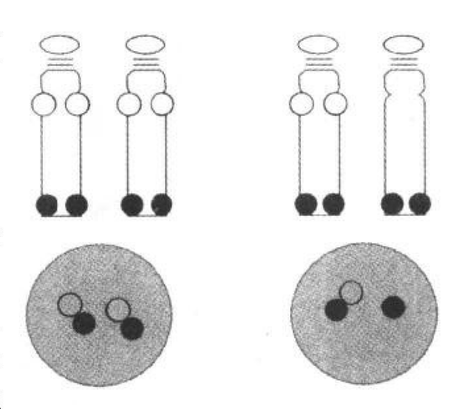


图 8-21 通过 FISH 分析检测 CATCH22 综合征的微缺失。在正常个体（左）和受累个体（右）的中期染色体和间期核中应用基因座特异性探针得到的结果。识别 22 号染色体长臂远端的对照探针（黑色）用来标记 22 号染色体。注意：在正常个体中，两个 22 号染色体和 22q11 探针（白色）有杂交信号，而在受累个体中，仅有一条 22 号染色体被标记（引自 Emanuel BS. The use of fluorescence in situ hybridization to identify human chromosomal anomalies. Growth Genet Horm 1993;9:6 ~ 12.）。

3. 基因组印刻

两个不同的微缺失综合征，Prader-Willi（PWS）综合征和 Angelman 综合征（AS）具有特别重要的意义，因为它们说明了一个新近才认识到的现象——**基因组印刻**（genomic imprinting），即依赖于父母亲源性的等位基因的差异性表达。PWS 在活产儿中的发病率为 1/10 000，它的特征是轻到中度的智力低下，婴儿期肌张力低下、反应差，身材矮小，手脚小，外生殖器小，巨食症（强迫性过食）和两岁以后的过度肥胖（图 8-22）。通过前中期染色体分析发现，大约一半的 PWS 患者有染色体 15q11 ~ q13 的微缺失。DNA 多态性分析很明显地表明，所有的 15 号染色体的缺失都是从父亲遗传的。AS 的表现不仅有更严重的智力低下、失语，而且还有癫痫、跳步、假笑、舌外伸和巨型下颌。令人奇怪的是，许多 AS 患者与 PWS 患者有相同的 15 号染色体缺失，但在这些 AS 患者中，有缺失的染色体通常是从母亲遗传的。这些现象说明某常染色体基因是差异性表达的，这种差异决定于该基因是来自父亲还是母亲。

通过 FISH 分析，现在已知道大约 70% ~ 75% 的 PWS 患者有父源性 15q11 ~ q13 的染色体微缺失。另外 20% ~ 25% 的患者通过多态性分析发现含有母源性 15 号染色体的两个拷贝而无父源性 15 号染色体的拷贝，这种情况称为**单亲双体性**（uniparental disomy）（图 8-23）。这种情况可能是由含有两条母源性染色体和一条父源性染色体的 15 三体性合子引起的。15 三体性通常是致命的，但是缺失一条 15 号染色体可能允许生存下去，产生可存活的婴儿。在这种三体性中，如果父源性染色体丢失，婴儿出生时就是母源性单亲双体性，将患 PWS 综合征。极少（ < 5%）的 AS 患者有 15 号染色体的父源性单亲双体性。这些自然事件表明，父亲和母亲对常染色体基因的遗传贡献不一定是相等的，而且对正常发育来说，来自父母双方的遗传贡献都是必不可少的。

更多基因组印刻的显著证据来自于早期鼠胚的研究。在鼠的卵子受精后，精子和卵子的原核在一段时间内仍保留分离状态，并能在显微镜下区分开来。在实验中，一个原核（假定是父源的）被除掉，另一个母源性原核插入这个卵子。具有两个母源核（双雌）的胚胎或多或少正常发育了一段时间，但是外胚膜发育严重不良。具有两个父源核（双雄）的胚胎虽有相对正常的外胚膜，但胚胎结构却严重异常。这两种情况没有一种与正常的胚胎发育相容，所以胚胎发育停止。对正常的发育来说，父亲和母亲的遗传贡献都是必需的，也是互补的。这与在父母配子形成过程中某些基因是差异性活化或失活（即印刻）

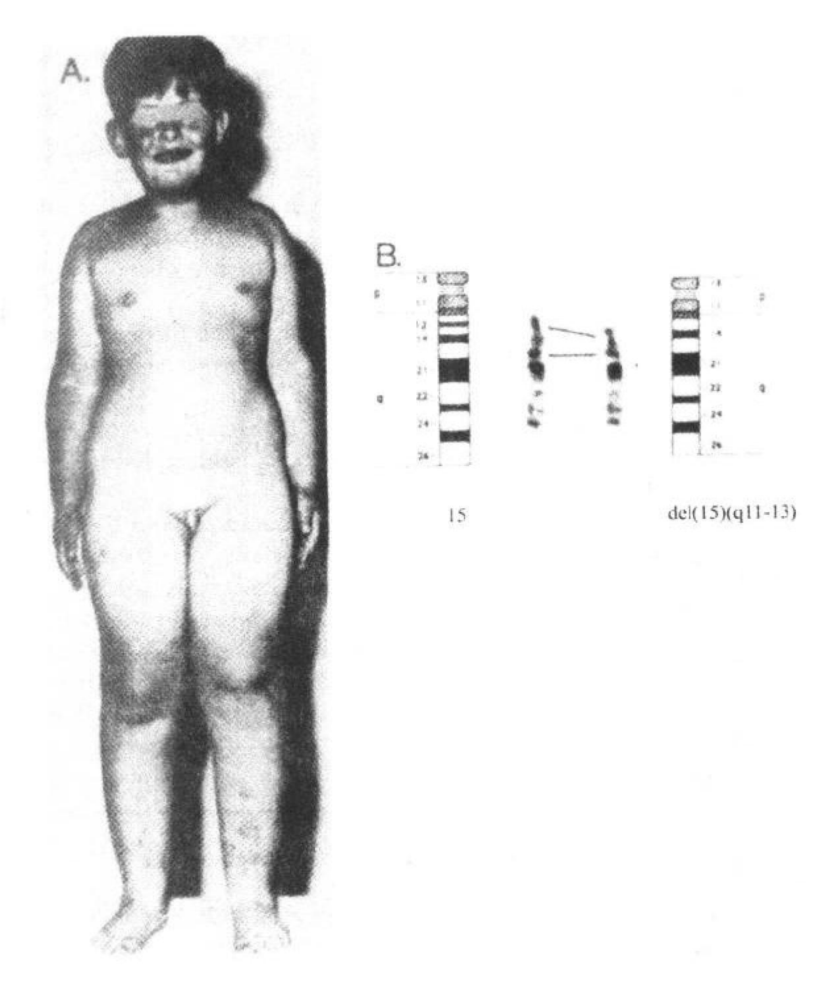


图 8-22 Prader-Willi 综合征。A. 一个 9 岁半的男孩，肥胖，阴茎小，手足小（引自 Jones KL. Smith's Recognizable Patterns of Human Malformation. 4th ed. Philadelphia: WB Saunders, 1988）。B. 一个 Prader-Willi 综合征病人的简略图和部分核型显示 15 号染色体长臂 11 到 13 带（15q11 ~ 13）的中间缺失（由 the University of Michiganf Clinical Cytogenetics Service 提供）。PWS 中间缺失的 FISH 诊断见图 8-8G。

这一观点一致。这样，在不同的基因座上，仅是母源或父源的基因能被表达。进一步的证据来自于两种人类疾病。葡萄胎是人类胎盘肿瘤，它有两套父源性单倍体染色体而无母源性染色体。卵巢畸胎瘤是一种含有多种组织的良性肿瘤，它由仅含母源性染色体而缺乏父源性染色体的细胞引起。这些疾病又一次表明父母中一方基因的缺乏会引起发育异常。

现在认为印刻的机制涉及在男性、女性配子形成过程中导致转录失活的基因的甲基化。因此，一个基因的活性等位基因（如 PWS 中的父源性等位基因）的缺失导致这一区域的结构性单体性和功能无效体，因为剩余的母源性等位基因不表达。携带失活等位基因（如 PWS 中的母源性二体性）同源染色体单亲双体性会导致功能无效体，尽管它在结构上是二倍体。基因组印刻可能有助于解释染色体缺失综合征表型的差异，这种差异依赖于

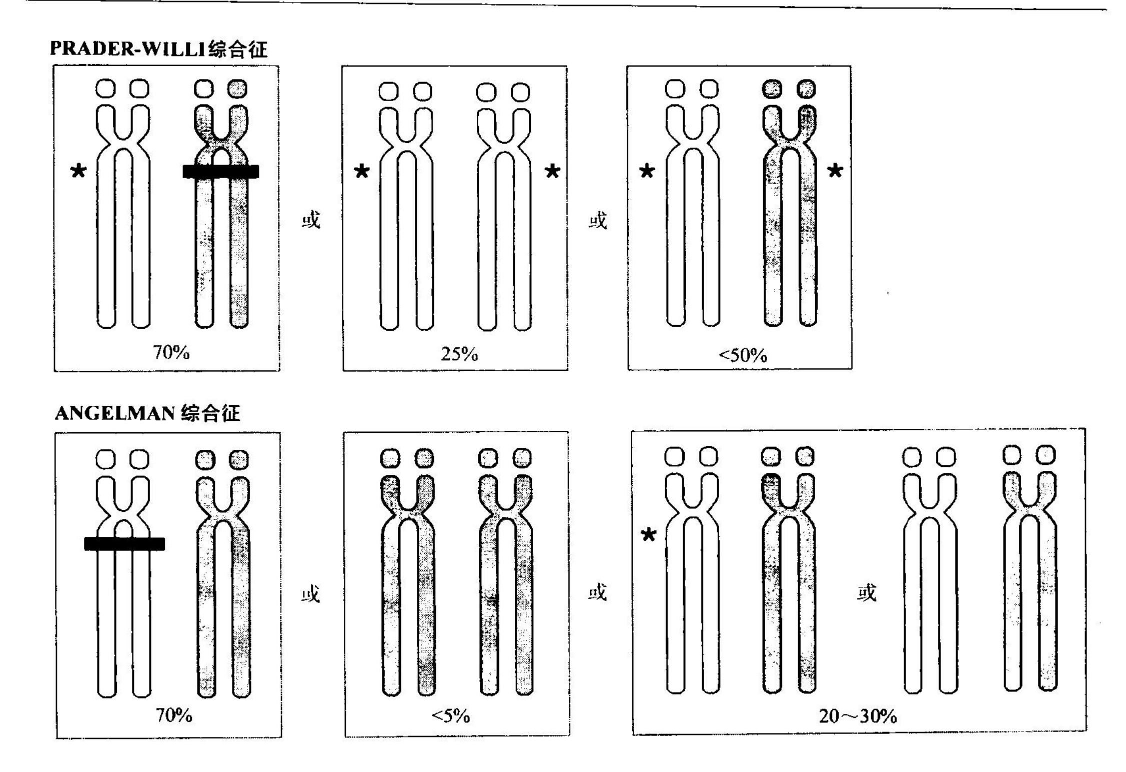


图 8-23 Prader-Willi 综合征(PWS)和 Angelman 综合征(AS)的分子基础。大约 70%到 75%的 PWS 患者有父源性 15 号染色体(灰色)部分的微缺失(黑杠)。在 15q11~13 的几个基因座处,母源性 15 号染色体(白色)甲基化。大约 20%到 25%的 PWS 患者有母源性单亲双体性,这两个母源性同源染色体甲基化,转录是静止的。一小部分患者有双亲遗传的两条 15 号染色体,没有可发现的缺失,但在两条 15 号染色体关键区域的几个基因座处表现出母源的甲基化。在 AS 中,70%的患者有母源性 15 号染色体的可发现的微缺失,但仅有不到 5%的患者有父源性单亲双体性。在家族性 AS 中,最常见的情况是遗传双亲源 15 号染色体,没有缺失,有正常的甲基化方式。人们推测这些病人在母源性 15 号染色体上有尚未鉴定的 AS 基因可能的点突变或其他微小的异常。最后,一些 AS 病人遗传双亲源 15 号染色体,伴有异常的甲基化(缺乏正常甲基化的母源等位基因)(由 D Ledbetter. National Center for Human Genome Research 提供图片)。

缺失的双亲来源。并且,对一些常染色体显性遗传病来说,它可能有助于解释不完全外显率和表型严重性的差异,这些差异依赖于突变基因的双亲来源。

(三) 肿瘤中的染色体改变

除了生殖系即先天性染色体异常之外,在多种实体瘤和白血病中对体细胞染色体异常的作用的认识正在不断发展。一个大家都清楚的特殊例子就是 Burkitt 淋巴瘤,该肿瘤通常有一个 8 号染色体长臂和 14 号染色体长臂的相互易位,少数为 8 号染色体长臂和 22 号染色体或 2 号染色体长臂的相互易位。在一类成人白血病——慢性粒细胞白血病中,几乎总有 9 号染色体长臂和 22 号染色体长臂之间的相互易位(Ph^1 或费城染色体)。这些易位引起恶性转化的机制将在第十一章中讨论。因为在多种白血病和淋巴瘤中都可见到特异性和一致性的染色体改变,所以包括 FISH 在内的染色体分析在它们的临床处

置中是一种重要的诊断和判断预后的工具。

除了以上讨论的肿瘤中的体细胞染色体异常外，还有一类罕见的隐性遗传疾病，它们以异常高频的自发性或诱导的染色体断裂以及白血病和实体瘤的高发病风险为特征。这些疾病包括第十一章讨论的 Bloom 综合征，Fanconi 贫血和共济失调性毛细血管扩张症，它们的基因定位和克隆将会提高我们对恶性转化的遗传基础的理解。

(四) 性染色体

"似乎在性别进化过程中，以前某个片段从 X 染色体上断裂下来……，然后就其他染色体而言，它不能阻碍这些染色体以一种非完全女性的形式自我表达。这种人我们称为男性。男性所继承的几乎所有麻烦都可以追溯到这种最初的男性 X 染色体缺陷。"

——ASHLEY MONTAGU（引自 Montagu Ashley. THE NATURAL SUPERIORITY OF WOMEN. MACMILLAN PUBLISHING COMPANY. © 1952, 1953, 1968, 1974 ASHLEY MONTAGU）

性染色体和性别分化

X 染色体很大，含有全部 DNA 的 6%。Y 染色体通常小得多，尽管它在大小上是变化的，这反映异染色质的不同含量。大约 250 种疾病是 X 连锁的，但仅有不到 20 种是 Y 连锁的。300 多个功能基因和 10 000 多个 DNA 标记已定位于 X 染色体，这些基因中大部分编码体细胞的功能。相比之下，虽然目前有 700 多个 DNA 标记定位于 Y 染色体，但仅有 26 个基因分配在 Y 染色体上。然而，其中一个基因称为 SRY，它在决定性腺性别上起关键性作用。

在胚胎生命早期，人类胎儿性腺是不分化的。在发育第四周，生殖细胞从卵黄囊移到性腺。大约在发育的 6～7 周时，性腺开始进行性别分化（图 8-24）。有 Y 染色体存在时，性腺的中央部分即髓质，发育成睾丸。睾丸合成两类诱导物，一类是睾丸酮，它引起 wolffian 管增殖。这些 wolffian 管发育成男性内生殖器。另一类是一种非固醇类糖蛋白，叫做 müllerian 管抑制物质，它引起 müllerian 管结构退化。在缺乏 Y 染色体时，如人类 XX 女性，性腺外层部分即皮质，发育成为卵巢。这时，没有睾丸酮的产生，也没有 Müllerian 管抑制物的产生。因此，wolffian 管结构退化而 müllerian 管结构增殖形成女性内生殖器，包括阴道上部、子宫和输卵管。通常是睾丸而不是卵巢的存在与否决定性别分化。在早期哺乳动物胚胎中除去性腺将引起女性内生殖器的发育。

从古代起，性别决定的机制就是科学假想和研究的主题。人们曾推测人类性别决定同果蝇（*Drosophila*）一样，由 X 染色体和常染色体数目的比率决定。直到 1959 年才认识到决定性别的是 Y 染色体存在与否。这是基于对表型为男性的 47, XXY Klinefelter 综合征个体的研究以及只有一条 X 染色体雌性小鼠的发现。

按照 Y 染色体决定性腺性别这一原则，通过对罕见特例的研究已寻找到 Y 染色体上负责睾丸发育的特异性基因。应用一系列定位于 Y 染色体不同区段的 DNA 探针，对具有 46, XX 染色体组成的男性进行研究。这些男性除了具有无精子发育的小睾丸之外，在

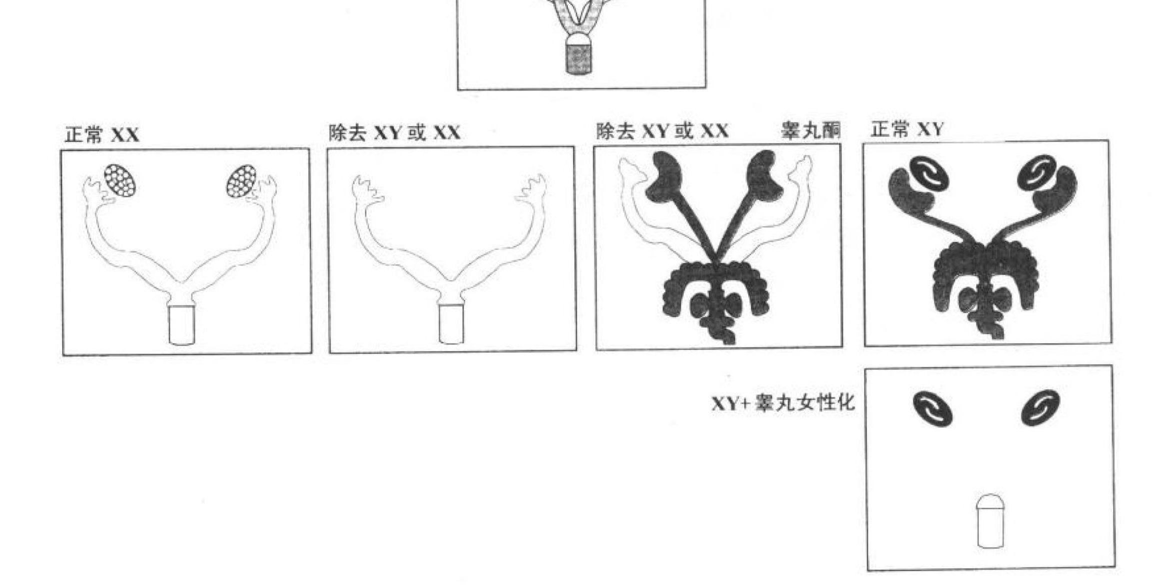


图 8-24 人类性别分化。详情见书。

表型上是正常的，这一点与以后要讨论的 Klinefelter 综合征一样。人们曾假设，由于 X 和 Y 染色体间易位，这些个体在父源性 X 染色体上有隐秘的 Y 染色体序列。和常染色体不同的是，在男性减数分裂中，X、Y 染色体并不随它们的长度进行肩并肩的排列，而是进行头对头的定向排列(图 8-25)。X、Y 染色体短臂末端含有一段高度同源的 2.6Mb 长的所谓假常染色体区域，在减数分裂时它们之间存在广泛的重组。靠近假常染色体区域的 Y 染色体短臂上的重组事件可能导致睾丸决定因子从 Y 染色体易位至 X 染色体上。事实上，在大约 85%的 XX 男性中发现含有 Y 来源的序列。Y 染色体上邻近假常染色体区域的一段 35kb 的区域被证明是引起 46,XX 个体性别逆转所必需要的、最少量的 Y 染色体 DNA(图 8-26)。对这一区域的测序鉴定了一个基因，它被命名为性别决定区域 Y，即 *SRY* (sex-determining Region Y)。这个基因含有一个高度保守的 DNA 结合域，与一类叫做高移动组(high mobility group, HMG)蛋白盒同源。已知这类蛋白可引起 DNA 弯曲和活化基因转录。分析鼠类这个基因的同源序列 *Sry* 的表达方式表明：正如所料，它是一个睾丸决定因子，只在睾丸即将分化前于性嵴的体细胞中表达。注意，按照规定，人类的基因符号都是大写的(如 *SRY*)，而啮齿类动物基因符号除第一个字身外都是小写的(如 *Sry*)。表明 *SRY*/*Sry* 事实上是睾丸决定因子的证据来自于对含有鼠 *Sry* 基因的 14kb 的片段的研究。该片段被用来制造含有正常雌性 XX 染色体以及携有 *Sry* 基因的 14kb 的 Y 染色体片段的转基因小鼠。这些在染色体上是雌性的小鼠有一些发育为具有睾丸和正常雄性交配行为的雄性表型。有趣的是，和人类 XX 男性一样，这些小鼠是不育的。

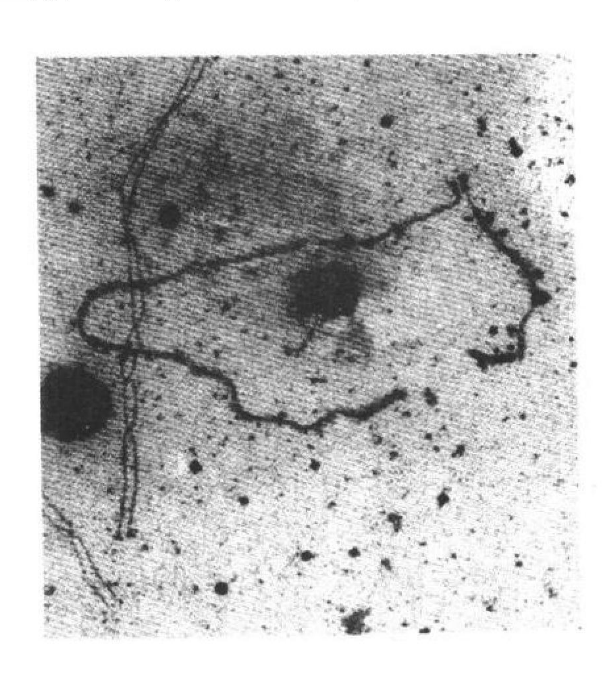


图 8-25 在减数分裂Ⅰ粗线期,X染色体(左)和Y染色体(右)的电镜照片。注意配对局限在两个染色体短臂末端假常染色体区域(引自 Connor JM. Ferguson-Smith MA. Essential Medical Genetics. 4th ed. Oxford: Blackwell Scientific Publications, 1993.)。

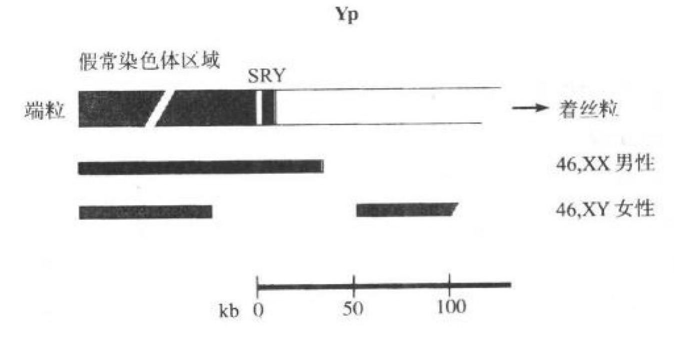


图 8-26 Y染色体性别决定区域图。假常染色区域以灰色表示而 *SRY* 基因以黑色表示。黑线表示在46,XX男性中发现的最小的Y染色体DNA量,灰线描绘在46,XY女性中缺失的区域(引自 Affara NA. Sex and the single Y. BioEssays 1991;13:475~478.)。

对46,XY女性的研究显示,一些人具有 *SRY* 基因的缺失,而另一些人有 *SRY* 基因的点突变和小缺失,事实上这些突变和小缺失都发生在高移动组蛋白盒的DNA结合区域。目前 *SRY* 基因内已有20多个突变被描述,但这些突变仅占XY女性的15%~20%。这说明睾丸决定途径中其他基因的突变也许能解释这种表型。这也和作为转录因子的 *SRY* 基因一致。*SRY* 必须激活睾丸分化途径上其他的基因才能使内生殖器进行正常发育。对其他罕见的性别逆转疾病的研究应有助于澄清这一途径中的一些步骤。

外部性别分化是一种引起男性外生殖器分化的雄激素衍生物(5α-双氢化睾丸酮)

影响的结果。在具有 X 连锁突变的睾丸雌性化个体中，存在一种异常或缺陷的雄激素受体，导致其对雄激素不敏感，所以这些个体不能发育正常的男性外生殖器。他们具有睾丸和一个正常女性外部表型(图 8-27)，但缺乏女性或男性内生殖器结构。当这些妇女不能正常行经时，她们在临床上被发现。研究显示核型为 46，XY，并通常在腹股沟区发现睾丸。



图 8-27 睾丸女性化病人。注意：除了腋毛和阴毛减少之外，具有正常的女性表型(引自 Grumbach MM，Van Wyk JJ. Disorders of sex differentiation. In：Wilson JD，Foster DW，eds. Textbook of Endocrinology. 7th ed. Philadelphia：WB Saunders，1985.)。

(五) X 染色体失活

就 X 染色体而言，尽管男性都是半合子(仅有一个拷贝)，但是 X 编码的基因产物如红细胞 G6PD 活性(图 8-28)的平均量与有两条 X 染色体的女性相同。因此，肯定存在某种剂量补偿机制。四位遗传学家 Mary Lyon、Lianne Russell、Ernest Beutler 和 Susumo Ohno 彼此独立的研究使这一机制得以理解。这种机制如今称为莱昂(Lyon)假说，以英国遗传学家 Mary Lyon 命名。莱昂假说(图 8-29)论述了在体细胞：(a)X 失活发生在胚胎生命**早期**(人类晚期囊胚期)；(b)失活是**随机**的，即父源或母源的 X 染色体失活；(c)X 失活是**完全**的，即事实上这条 X 染色体全部失活；以及(d)X 失活是**永久的**和**克隆式繁殖**的，即在某个特定的细胞中，如果父源性 X 染色体失活，那么所有这个细胞的子代细胞中都将表达有活性的母源 X 染色体而父源 X 染色体失活。莱昂化的结果造成女性都是细胞的**嵌合体**(图 8-30)，每一个都是这条或那条 X 染色体功能上的半合子。

应当注意的是，以上关于 X 失活的陈述也存在例外。虽然 X 失活通常是随机的，但一个结构异常的 X 染色体，如一个有缺失的 X 染色体是优先失活的。另一方面，在携带 X 常染色体平衡易位的个体中，通常是正常的 X 染色体优先失活(见第九章)。虽然 X 失活是广泛的，但它不是完全的。已知有 16 个以上的基因逃避了这种失活(图 8-31)。这些基因包括位于或邻近假常染色体区域的基因以及远离这一区域的基因。这些基因中有一些但不是全部，具有功能上的 Y 同源物，它们可能维持性别之间的剂量相等。据估计，人类 X 染色体上多达 1/3 的基因可能逃避完全失活。最后应注意，尽管在大多数体细胞中 X 失活是永久的，但在生殖细胞发育中，这一失活必须是可逆转的。

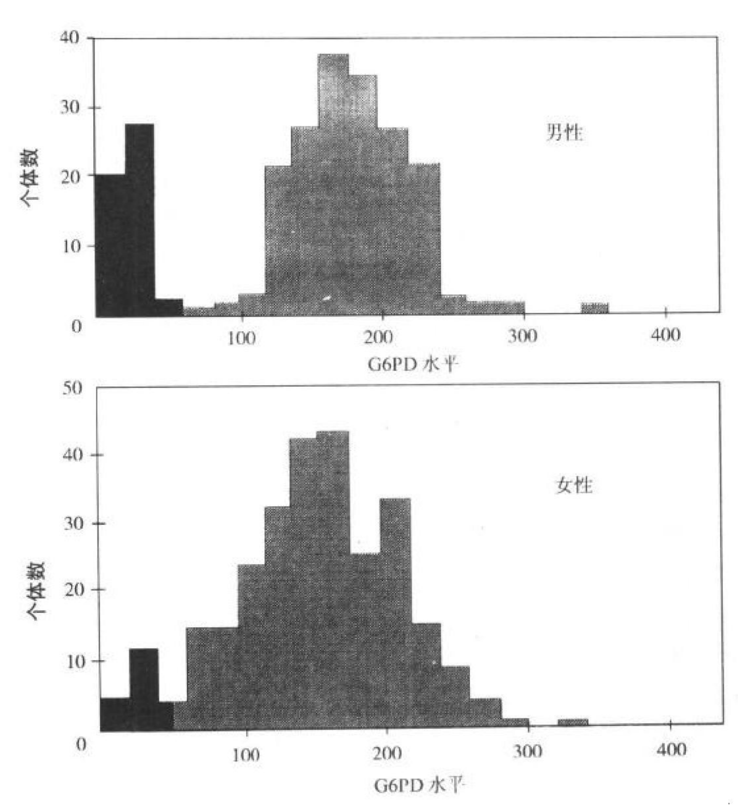


图 8-28 在从尼日利亚抽样的黑人男性和女性人群中红细胞 G6PD 活性分布。黑色栏表示 G6PD 缺陷个体。注意：非缺陷个体（灰色阴影区）G6PD 活性的分布在男性和女性相同（引自 Harris H. The Principles of Human Biochemical Genetics. 2nd ed. Amsterdam: North-Holland Publishers, 1975.）。

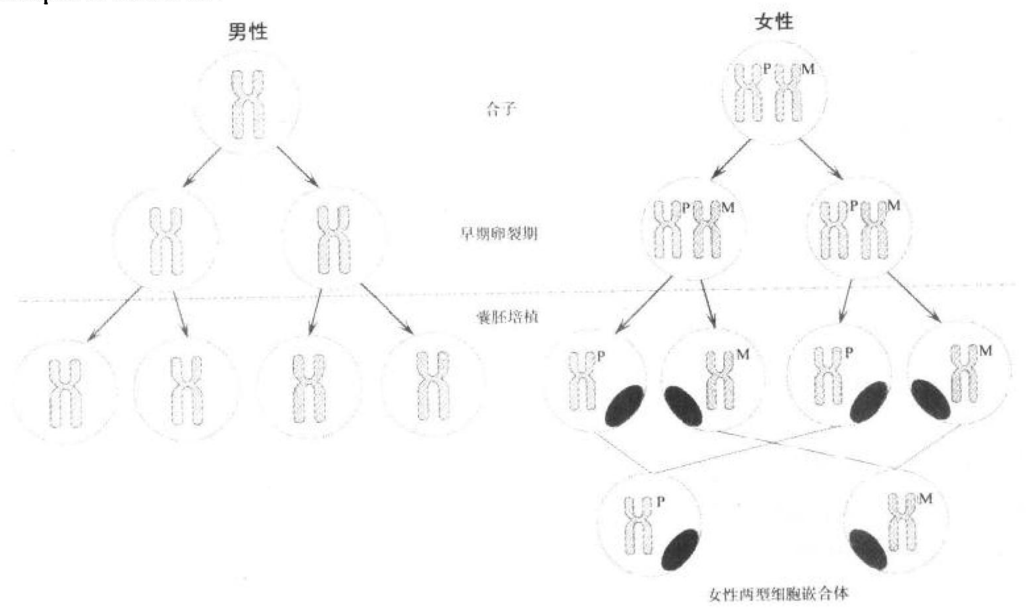


图 8-29 莱昂假说的简图。在女性中，*P* 代表父源 X 染色体，*M* 代表母源 X 染色体。男性惟一的 X 染色体肯定是从母亲那里遗传的。失活 X 以黑色团块表示。

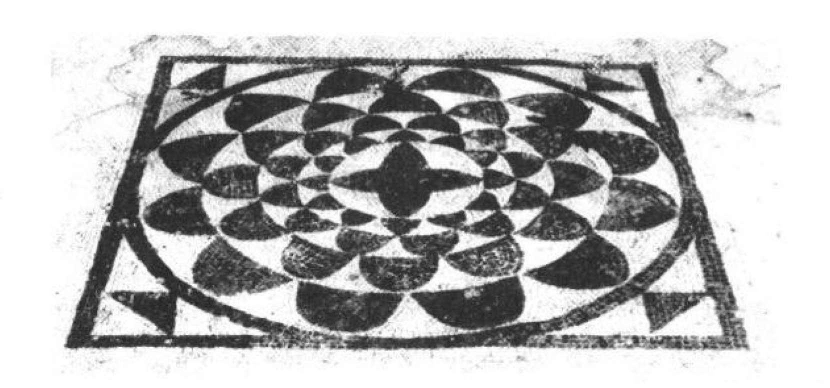


图 8-30 罗马 Prompeii 别墅地板的马赛克。注意仅是黑色石头的图块和仅为白色的图块。

几方面的证据支持莱昂假说。第一个证据来自于小鼠的遗传研究。一种 X 连锁的毛皮颜色突变(如:玳瑁色)的雌鼠杂合子通常被发现有斑点状皮毛,即它们呈现玳瑁色皮毛和野生型皮毛混合的斑块(图 8-32)。相比之下,这种突变的雄性半合子则呈现统一的玳瑁色皮毛。花纹猫是这一现象的另一个相似的例子。雌鼠是两个非等位基因的杂合子,当这两个基因是互斥的时候(在相反的染色体上),突变的毛色基因表达呈现含有这种或那种突变颜色的皮毛;而当这两个基因是偶联的时候(在同一条染色体上),皮毛斑点含有混和的两种突变颜色或呈现野生型颜色。

莱昂化的细胞学证据首先来自于性染色质,即 Barr 小体的意外发现。加拿大细胞学家缪瑞·巴尔在研究反复电刺激对猫神经元细胞产生的形态学作用时注意到,在一些猫的神经元的核中含有一个额外的深染的染色质小体,这一染色质小体仅在雌猫出现。接下来利用从颊黏膜刮下来的上皮细胞对人类进行的研究表明,性染色质的个数等于 X 染色体数减 1。正常男性只有一条 X 染色体,无 X 染色质;而正常女性含有两条 X 染色体,有一个 X 染色质(图 8-33)。有三条 X 染色体的个体有两个 Barr 小体。后来的研究显示,在前期,两条 X 染色体中一条呈现异固缩,即致密而深染,暗示为失活的染色质。最后,对培养细胞应用胸腺嘧啶标记染色体的研究表明,两条 X 染色体中有一条在细胞周期的 S 期的晚

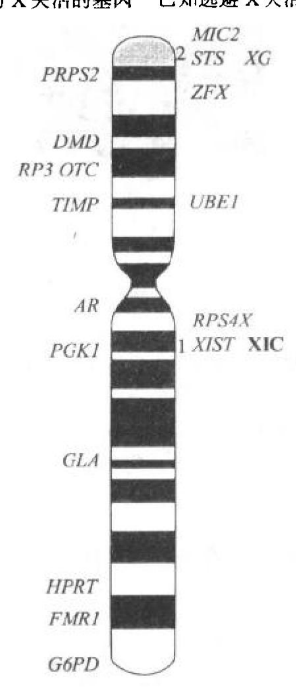


图 8-31 描绘经历 X 失活和其他逃避 X 失活的基因的人类 X 染色体部分图。*DMD*,抗肌萎缩蛋白基因在 Duchenne 肌营养不良中是异常的;*OTC*,鸟氨酸氨基甲酰转移酶基因;*AR*,雄激素受体;*HPRT*,次黄嘌呤磷酸核糖转移酶基因;*G6PD*,葡萄糖-6-磷酸脱氢酶基因;编码细胞表面蛋白的 *MIC2* 和 *XG*;*STS*,类固醇硫酸酯酶基因;*ZFX*,锌指 X 基因;*RPS4X*,核糖体亚单位蛋白 4 基因;*XIST*,X 失活特异性转录物。X 失活中心(XIC)在 Xq 上是 1 带;假常染色体区域是 Xpter 处 2 区域(引自 Davies K: The essence of inactivity. Nature 1991;349:15~16.)。

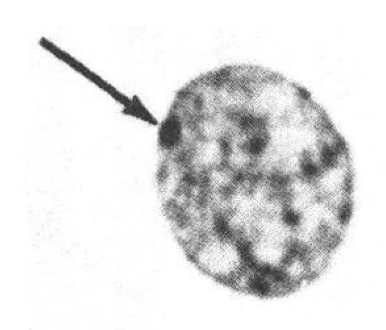
期进行复制。

支持莱昂化的生化学证据首先来自对 G6PD 杂合性妇女培养的皮肤成纤维细胞中 G6PD 表达的研究。这种妇女具有 G6PD 常见的电泳变异体 G6PDA 和野生型 G6PDB 等位基因。大量培养的成纤维细胞表达这种酶的 A、B 两种形式，而皮肤成纤维细胞克隆，即来源于单一前体细胞的细胞系，仅表达 A 或 B，而不是两者都表达(图 8-24)。类似的具有克隆表达的生化标记，包括鸟苷酸氨基甲酰转移酶已被确定为 X 连锁的基因(图 8-35)。生化学证据和细胞学证据之间的相关性来自于对骡，即一种马和驴杂交产生的雌性后代的研究。因为马和驴的 X 染色体可以区分开来，G6PD 活性差异可以通过电泳进行分辨，因此，在骡的皮肤成纤维细胞单个细胞系中，驴的 X 染色体失活时出现马的 G6PD，反之亦然。因此，我们得出相关性的结论。遗传学、细胞学以及生化学证据综合起来，有力地支持了在雌性哺乳动物的每一个体细胞中只有一条 X 染色体具有活性这一假说。



图 8-32 X 连锁皮毛颜色突变的雌鼠玳瑁色杂合子(引自 Thompson MW. Genetic consequences of heteropyknosis of an Xchromosome. Can J Genet Cytol 1965; 7: 202 ~ 213.)。

莱昂假说的重要性是多方面的。首先，它有助于理解人类 X 连锁的基因和疾病的多样性，并且它能解释为什么杂合性女性在临床表现上比半合子的男性具有更大的可复性。同样，它有助于解释为什么鉴定 X 连锁疾病的女性携带者的生化携带者识别技术存在相当的困难。应当注意，尽管 X 失活通常是随机的，但在个体某一特定的



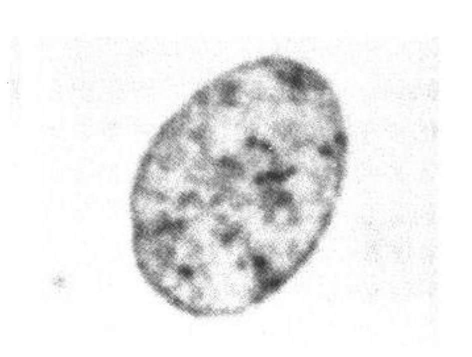


图 8-33 人类颊黏膜细胞的性染色质小体。注意 Barr 小体(箭头)在女性中出现(左)但在男性中缺如(右)(引自 Grumbach MM, Van Wyk JJ. Disorders of sex differentiation. In: Wilson JD, Foster DW, eds. Textbook of Endocrinology. 7th ed. Philadelphia: WB Saunders, 1985.)。

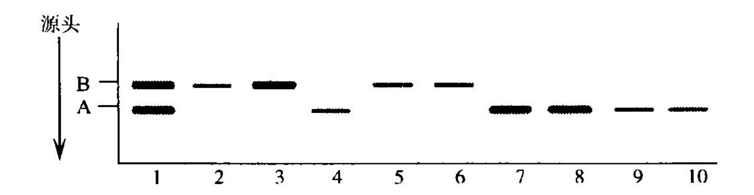


图 8-34　一个 G6PD 电泳变异体 A 的杂合性妇女成纤维细胞培养的 G6PD 电泳分析。注意在成纤维细胞群中(第 1 道),G6PDA 和 G6PDB 都表达,而在成纤维细胞分离克隆群中仅表达 G6PDA 或 B,从不两者都表达(引自 Davidson RG. Nitowsky HN. Childs B. Demonstration of two populations of cells in the human female heterozygons for glucose-6-phosphate dehydrogenase varianсs. Proc Natl Acad Sci USA 1963;50:481 ~ 485.)。

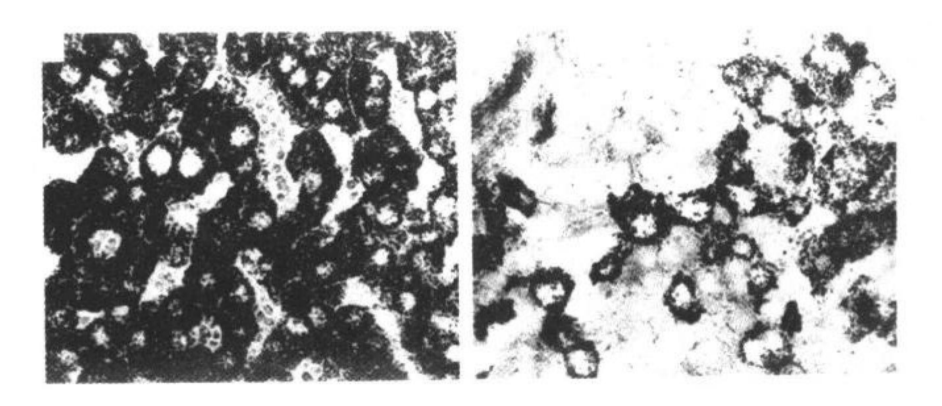


图 8-35　一个正常妇女(左)和一个 OTC 缺陷的杂合性妇女(右)的肝活检组织的 OTC 活性的组化染色。肝活检组织经生化测定后,后者有 21% 的正常 OTC 活性。然而,单个细胞或有正常的活性(深染)或缺乏活性(引自 Ricciuti FC, Gelehrter TD, Rosenberg LE: X-Chromosome inactivation in human liver: confirmation of X-linkange or ornithine transcarbamylase. Am J Hum Genet 1976;28:332 ~ 338.)。

细胞谱系中,也可能发现明显弯曲的 X 失活方式。X 染色体失活作为一种标记,还被用于研究分化和恶性转化。比如,子宫纤维肌瘤(一种常见的良性子宫平滑肌瘤)的研究表明,在 G6PDA 和 B 的杂合子女性中,每一个子宫纤维肌瘤仅表达 A 型或 B 型的 G6PD,从不同时表达这两种。这被解释为提示说明每一个肿瘤最有可能起源于单个细胞。X 连锁 DNA 标记的使用明显提高了这一技术的作用。最重要的可能是,X 失活为基因调节的研究提供了一个有趣的模式,因为它提供了一个在同一条染色体上的大多基因都是永久失活的例证。

尽管 X 失活的机制还没有得到完全理解,但它肯定涉及染色质结构的改变,包括 DNA 的差异性甲基化和组蛋白乙酰化。在失活 X 染色体上,沉默基因的 CpG 岛(在基因 5′上游区域内的 CpG 二核苷酸簇)是高度甲基化的,而它们在有活性的 X 染色体上的同源序列则是非甲基化的。在失活的 X 上逃避了失活的基因的 CpG 岛仍保持非甲基化。对改变了 X 失活方式的 X 染色体易位和缺失的遗传学分析已经在 Xq13.2(图 8-31)鉴定了一个 X 失活中心(X-inactivation center, XIC)。对 X 失活发生来说,这个中心要求必须是

顺式的。在 XIC 内有一个候选基因——X 失活特异性转录物(*X Inactive-Specific Transcript*, *XIST*,读作 exist),它只在失活的而非有活性的那条 X 染色体上表达。因此,在核型正常的个体中,它的表达是女性特异性的;但在 X 染色体非整倍体的个体中,*XIST* 转录物与 X 染色体数减 1 成比例。通过同源重组造成的 *Xist* 基因(*XIST* 的鼠类同源基因)失活完全阻止了靶染色体的失活,说明 *Xist*/*XIST* 基因对 X 失活来说是必不可少的。*XIST* 基因转录生成一个 17kb 的 mRNA,它不含可读框(即含有许多终止密码子)。因此,*XIST* 基因的产物似乎是一个不被翻译的 RNA,它可能起一种结构性作用,而不是编码某一蛋白质。有趣的是,发现 *XIST* 转录物与间期核中的 Barr 小体有关。*XIST* 启动和/或维持 X 失活的机制仍属未知。

(六) 性染色体疾病

1. Klinefelter 综合征

Klinefelter 综合征,大约每 1 000 个活产男婴中就有一个发病,它以青春期后睾丸发育不良为特征。病人表型为男性,小睾丸,睾丸小管玻璃样变和精子缺乏(不能产生正常数量的精子),导致不育和不同表现的性腺发育不足。这些个体在青春期前表现正常,但在青春期后,常由于小睾丸,偶尔的男性巨乳和身体比例异常而被检查出来(图 8-36)。有趣的是,患有这种综合征的个体具有比预期要高的社会病理学发病率,他的表现可以从缺乏正常的社会调整心理到认识法律的实际困难。尽管他们的平均 IQ 有所降低,但明显的智力低下通常不在该综合征的表现之内。典型的 Klinefelter 综合征患者染色体组成为 47,XXY,他们的细胞呈现染色质阳性(也就是,他们含有一个 Barr 小体)。在他们一些细胞中至少含有两条 X 染色体以及一条 Y 染色体。这种形式的 X 染色体非整倍体被认为是由减数分裂不分离所引起的,不分离发生在父方和母方的机率均等。父亲的不分离事件(肯定发生在减数分裂Ⅰ期)和母亲的减数分裂Ⅰ不分离显示出一种双亲的年龄效应。一些 Klinefelter 综合征男性含有两个以上的 X 染色体拷贝和一个 Y 染色体,这说明发生了不止一次的不分离事件。患者含有的 X 染色体数越多,出现与该病有关的智力低下的可能性就越大。

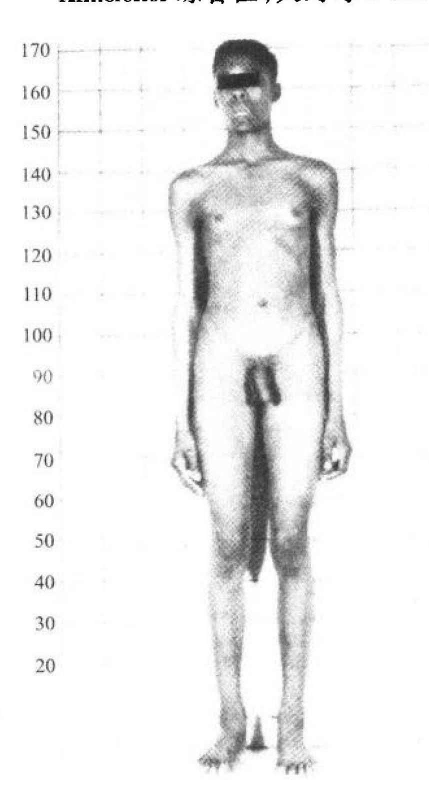


图 8-36 Klinefelter 综合征。注意:患者有一正常的男性生殖器但也有男性女乳(类似女性的乳房发育)(引自 J Chronic Dis (July 1960) Feinstein AR, ed., and in Medical Genetics 1958 ~ 1960. St. Louis: Mosby, 1961. Reprinted with permission of the editors of the Journal of Chronic Diseases.)。

2. XYY 综合征

由于 Klinefelter 综合征与社会病理学有关，所以在英国的精神刑事医院内进行了一些调查研究，用来估计在这些医院的犯人中 47,XXY 核型的频率。这导致 Y 染色体非整倍体的发现，如 47,XYY 和 48,XXYY。大约每 1 000 个活产男婴中就有一个 47,XYY 患者，但在这些精神刑事医院的犯人中，它的发现率为 4/1 000～20/1 000。与 Klinefelter 综合征表型有关的 48,XXYY 在这些犯人中的频率是新生儿人群的 50 倍。这就出现了一个问题——多余的 Y 染色体是否通常与行为异常有关。XYY 男性有正常的性发育和生育能力。惟一与 XYY 核型有关的表型是身材高大。与这种综合征有关的社会病理学风险级别(如果有的话)仍不清楚。Y 染色体的非整倍体肯定是父亲减数分裂不分离引起的。

3. Turner 综合征

Turner 综合征以性腺发育不全和性不成熟的女性表型为特征。在特征上，这些妇女有原发性闭经(不能行经)和不孕。典型的 Turner 综合征还有身材矮小(通常低于 5 英尺)和一些异常体征，包括颈蹼，婴儿期手足肿胀，肘提携角增大和心血管以及肾脏畸形(图 8-37)。活产女婴中 Turner 综合征发病率为 1/5 000，然而在胎儿中的 X 单体性的频率高达 4%。据估计，99%以上的 45,X 胎儿都自发流产了。

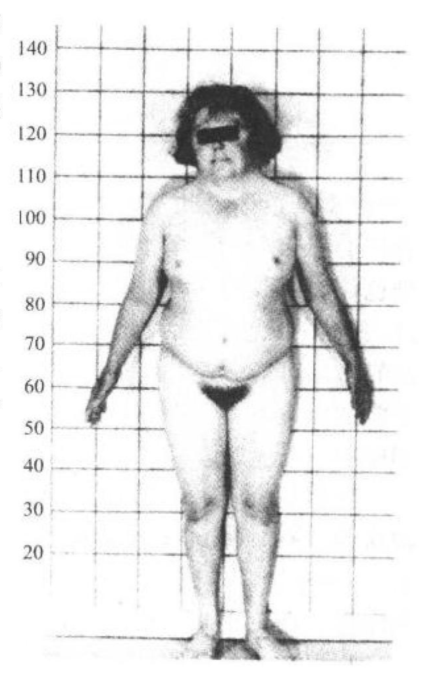


图 8-37 Turner 综合征。患者身材矮小，有颈蹼，胸廓膨隆伴肋间隙扩大，乳房未发育以及手臂提携角增大(引自 J Chronic Dis (July 1960) Feinstein AR. Ed. And in Medical Genetics 1958～1960. St. Louis: Mosby, 1961. Reprinted with permission of the editors of the Journal of Chronic Diseases.)。

在核型上，至少在某些 Turner 综合征患者的部分细胞中，第二条性染色体也有毛病(表 8-4)。略多于一半的 Turner 综合征患者是 X 单体性，即 45,X，X 染色质是阴性的。大约 15%有 X 染色体的结构异常，包括等臂 X 染色体、X 染色体部分缺失或环状 X 染色体。这些个体有 46 条染色体，X 染色质阳性，即有 Barr 小体存在。还有 15%是 45,X/46,XX 的嵌合体，另外 15%是 45,X 细胞系和 46,X，异常 X 细胞系或 47,XXX 细胞系的嵌合体。嵌合体的出现表明不分离发生在受精后的有丝分裂而不是在减数分裂。因此，Turner 综合征可由减数分裂不分离(在大约 80%的病人中，这种不分离发生在父亲的减数分裂)引起，也可由有丝分裂不分离引起，后者导致嵌合体产生。

从对含有一条异常 X 染色体的 Turner 综合征的研究中，人们提出 X 染色体短臂的单体性是引起身材矮小和伴随这种综合征出现的异常体征的原因，而 X 染色体长臂上的基因是卵巢发育和维持的关键。因为在女性细胞中只有一条 X 染色体是有活性的，因此，为什么会发生Turner综合征这一问题就提出来了。看起来似乎是位于失活X染色体上

表 8-4 Turner 综合征的异常核型

45,X		
46,X,异常 X		
缺失	X,Xp -	
等臂染色体	X,i(Xq)	短臂单体 长臂三体
环状染色体		
嵌合体		
X/XX		
X/XY		女性化,男性女乳
X/X,异常 X		
X/XXX		
X/XX/XXX 等		

(和在 Y 染色体上,因为 XY 个体无 Turner 综合征)的一个或多个基因能够阻止 Turner 综合征的临床表现。多年前,Ferguson-Smith 猜想引起 Turner 综合征的基因是那些逃避了 X 染色体失活并有 Y 连锁同源基因的 X 连锁基因。*RPS4* 基因编码核糖体蛋白亚单位,它逃避了 X 染色体失活,并在 Y 染色体上有同源基因,它作为解释 Turner 综合征的一些临床异常的可能候选基因而被提出。也有可能在胚胎发育早期,两个有活性的 X 染色体是正常的卵巢发育所必需的。已知在胚胎发育的减数分裂开始时,失活的 X 染色体在卵原细胞中重新激活。

卵巢是个不一般的器官,因为它可程序化破坏自身的主要产物——卵子。胎儿卵巢被认为含有 7 000 000 个卵母细胞。到出生时,卵母细胞数降至 3 000 000。到月经初潮时,卵母细胞数下降到大约 400 000,而到绝经期时,大约为 10 000。Turner 综合征个体在胚胎期有卵母细胞,但这些卵母细胞到两岁时都消失了,绝经在月经初潮之前就发生了。

根据对 Turner 综合征嵌合体患者的研究发现,正常的 46,XX 细胞系的存在减轻了该病的表现。有极少数的 46,XX/45,X 嵌合体患者有正常的月经,甚至有生育能力。XY 细胞系的存在(在大约 5% 的患者中发现)可能与出生时以及青春期的男性化有关,并与退化性腺的 20% 的恶性化风险具有更为明显的关系。有意思的是,46,XY/45,X 的 Turner 综合征嵌合体妇女是由男性合子引起的,该合子发生了一次有丝分裂不分离而导致了 Turner 综合征。

4. X 连锁智力低下和脆性 X 染色体

当外周围血淋巴细胞培养基缺乏叶酸或胸腺嘧啶代谢紊乱后,在染色体上就可以观察到明显的断裂或间隙,这些断裂或间隙称为脆性部位。运用这些技术,在患有一种 X 连锁的智力低下的男性发现了一个叶酸敏感的脆性位点。这个 X 连锁的智力低下就是后来知道的脆性 X 综合征。脆性 X 综合征的特征是面部异常,包括脸部狭长伴有突出的前额、下颌和耳朵(图 8-38),90% 青春期后的患者出现巨睾,轻度的结缔组织异常以及更重要的中到重度智力低下。受累女性的表现通常要轻。事实上,在细胞遗传学水平,在所有的临床受累男性和大部分受累女性的中期染色体涂片中,有 50% 可以观察到一个位

于 Xq27.3 的脆性位点，称为 FRAXA(*FRA*gile site, *X* chromosome, *A* site)(图 8-39)。脆性 X 染色体综合征的发病率在男性为 1/1 250，在女性为 1/2 000，并且各种族人群都可受累。因此，脆性 X 染色体综合征是人类最常见的一种遗传性智力低下。

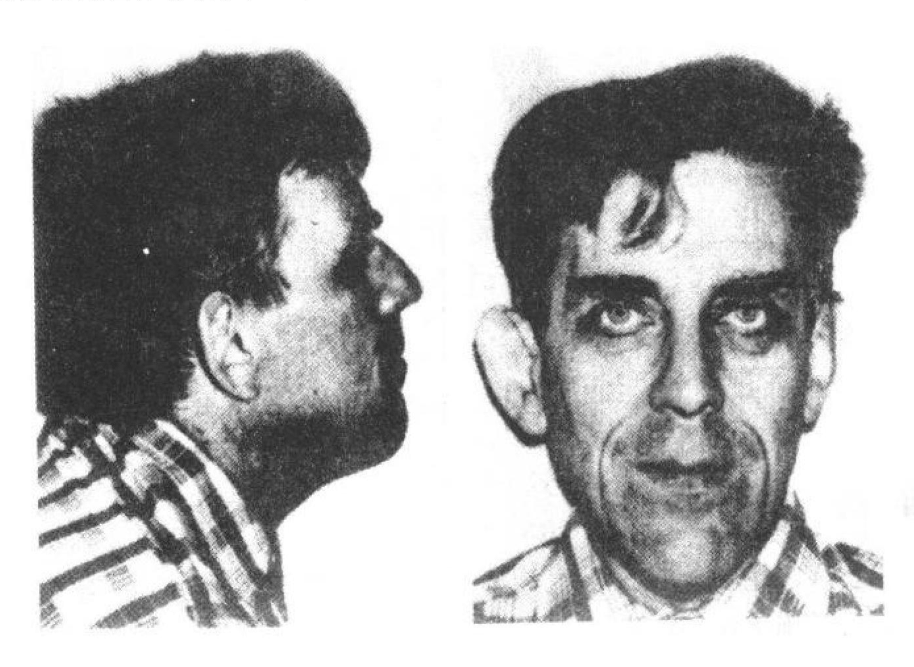


图 8-38 45 岁脆性 X 染色体综合征的男性。注意长脸伴耳、下颌和前额突出(引自 Jacobs PA, Glover TW, Mayer M, Fox P, Gerrard JW, Dunn HG, Herbst DS. X-Linked mental retardation: a study of 7 families. Am J Med Genet 1980;7:471 ~ 489.)。



图 8-39 四个智力低下和脆性 X 综合征的男性的外周血淋巴细胞培养的部分核型(仅显示 X 染色体)(引自 Jacobs PA, Glover TW, Mayer M, Fox P, Gerrard JW, Dunn HG, Herbst DS. X-Linked mental retardation: a study of 7 families. Am J Med Genet 1980;7:471 ~ 489.)。

尽管脆性X染色体综合征的遗传方式很明显是X连锁的，但是它仍有一些特点不同于单纯的X连锁遗传。大部分携带突变的男性有智力低下并表现出细胞遗传学异常，但全部男性肯定携带者的20%在表型上和细胞遗传学上都是正常的，称为正常的传递男性(transmitting male)。另一方面，1/3的女性杂合子临床上受累，但只有突变是遗传自母亲而非遗传自正常传递父亲时，她们才出现智力低下。引人注意的是，正常传递男性的孙子(女)智力低下的发病风险要高于其子女。因此，这种疾病的外显率或临床表现风险，在家系中似乎具有位置功能性，并且在连续后代中显示早发趋势。脆性X综合征这种不同寻常的系谱特点被称为**Sherman悖论**(图8-40)。

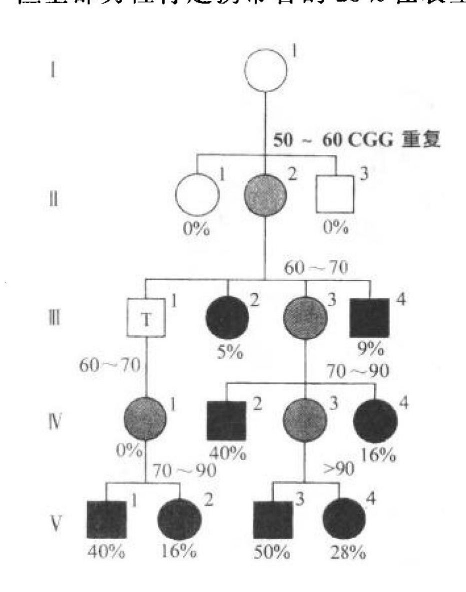


图8-40　解释Sherman悖论的假想家系。黑色符号表示智力低下的个体，浅灰色符号是临床未受累的携带者。T表示一个正常的传递男性。每个符号下的百分数字表示出现智力低下的风险，另一些数字表示发生这些风险的CGG重复数目(引自Fu YH, Kuhl DPA, Pizzuti A, Peiretti M, Sutcliffe JS, Richards S, Verkerk AJMH, Holden JJA, Fenwick RG, Warren ST, Oostra BA, Nelson DL, Caskey CT. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 1991;67:1047～1058.)。

对这种悖论的解释来自于命名为*FMR1*(*Fragile X Mental Retardation-1*)基因的克隆和描述。*FMR1*基因，即脆性X染色体智力低下基因，位于Xq27.3，长38kb，包含17个外显子，可能编码一种RNA结合蛋白。在非翻译的第一外显子内发现了一个CGG三核苷酸的重复序列，正常情况下其长度为6～50个重复次数，平均为30(图8-41)。在脆性X综合征受累患者中，这种CGG重复极度扩展至几百甚至几千次。当重复次数超过大约230时，*FMR1*基因5′端高度甲基化，转录关闭。*FMR1* mRNA和蛋白质的缺乏被认为是造成这种综合征临床表现的原因。这种扩展后的重复在有丝分裂中也是高度不稳定的。受累个体表现广泛的体细胞嵌合性，与不同组织中的重复次数不同有关。具有这种完全突变的所有男性和大约50%具有这种突变的杂合女性一样，在临床上发病。可以推测，女性临床上发病的可能性反映了X失活的方式。

大多数脆性X综合征的女性携带者具有少量数目的CGG重复，从大约50到230不等。在此范围内，*FMR1*基因5′区域是非甲基化的，基因具有转录活性。这种情况称为前突变(premutation)，可发现于正常传递男性中。前突变在传递过程中是不稳定的，并且由于某种未知的原因，当前突变是从母亲传递过来时，重复次数的增加要比突变从父亲传递过来时更为显著。在母亲减数分裂时，从前突变扩展到全突变(full mutation)的风险与重复次数成比例。当重复次数小于70时，风险低于20%，而当重复次数大于90时，风险大于99%。因此，全突变不是起源于正常长度的重复，而仅来自于前突变。受累女性是前突变母亲的女儿，而不是前突变父亲的女儿。因为三联体的重复次数似乎在家系传递过程中趋于增加，所以这一临床综合征的外显率在后代中似乎也是增加的。因此，这种三联体的动态重复解释了Sherman悖论。

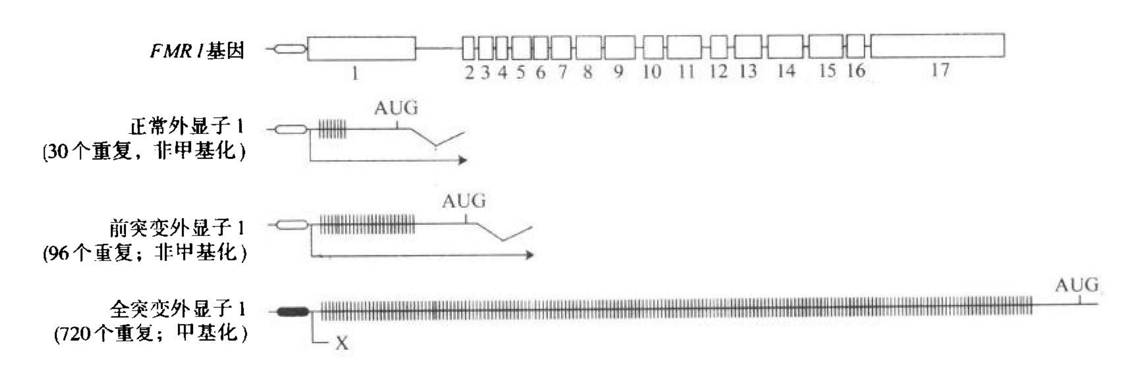


图 8-41 *FMR1* 基因和正常、前突变和全突变等位基因的第一外显子的图解。转录起始位置左边的椭圆形代表 *FMR1* 基因的启动子区域。空白符号代表活性转录，黑色符号代表转录静止。垂直线表示在转录起始点蛋氨酸密码子(AUG)上游的 CGG 三核苷酸(引自 Warren ST. Nelson DL. Advances in molecular analysis of fragile Xsyndrome. JAMA 1994; 271: 536～542.)。

任何一个未诊断的发育迟缓或智力低下的孩子都应该考虑是否为脆性 X 综合征。既然已经理解了这一疾病的分子基础，那么，分子诊断技术是可用的，它比细胞遗传学分析要更为有效和可靠。检查全突变的较好的方法是对产前或出生后获取的血液和组织的 DNA 进行 DNA 印迹分析。DNA 通常用两种限制性酶进行消化，其中一种酶对 DNA 甲基化状态敏感，这样就可以进行甲基化分析并能对三联体的大小进行估计(图 8-42)。

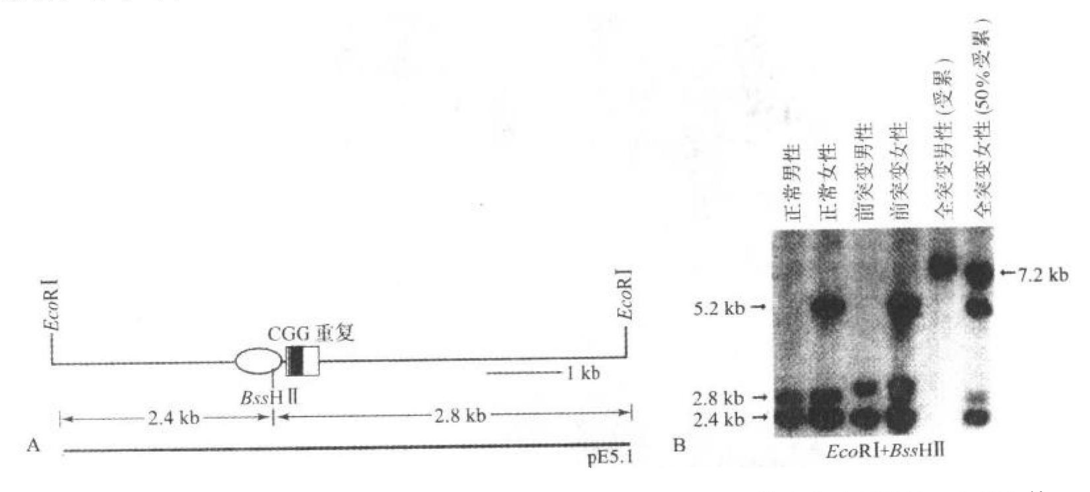


图 8-42 *FMR1* 基因 DNA 印迹分析。左图为 *FMR1* 基因简图。椭圆形代表启动子区域。方框为第一外显子，三核苷酸重复区域以黑色表示。*Eco*RⅠ片段正常为 5.2kb，*Bss*HⅡ活性对 DNA 甲基化状态敏感，仅能切割未甲基化的 DNA。右图为用 *Eco*RⅠ和 *Bss*HⅡ切割并与 pE5.1 探针杂交的 DNA DNA 印迹分析例子。在失活 X(在这条染色性体上，*FMR1* 通常是甲基化的)可观察到一条 5.2kb 带，而在女性的活性 X 染色体和正常男性的单一 X 染色体上可观察到两个较小的带(2.8kb 和 2.4kb)。从图中可见，CGG 重复位于 2.8kb 片段内，因此前突变男性和女性出现略大于正常 2.8kb 的带。全突变男性出现明显增大的甲基化带，而全突变女性出现三条带，即正常女性的两条带加一条额外增大的条带，它反映了甲基化和 *FMR1* 基因动态突变(引自 Warren ST, Nelson DL. Advances in molecular analysis of fragile X syndrome. JAMA 1994;271:536～542.)。

然而,对正常的或前突变的等位基因重复次数的精确判定最好应用 PCR 技术(图 8-43)。和本书讨论的许多其他遗传疾病不同,脆性 X 综合征是不典型的,因为绝大部分患者有明确的突变类型,即在 *FMR1* 基因 5′非翻译第一外显子这一特定区域内的三联体动态重复。这就极大地简化了它的分子诊断实验,但仍应该进行常规的细胞遗传学分析以排除可以解释患者临床特征的其他染色体异常。由于脆性 X 综合征的高发病率,所以正在考

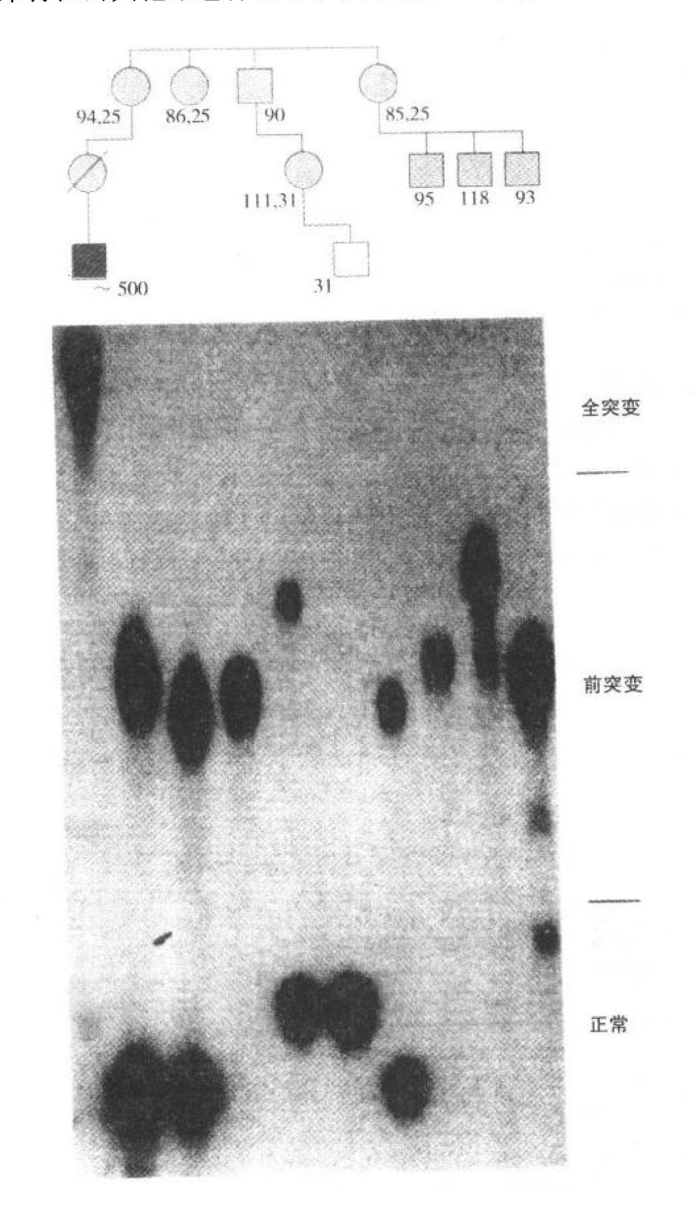


图 8-43 分离前突变的一个家族中个体的 DNA 的聚合酶链式反应分析。PCR 产物通过与放射性标记的 CGG 寡核苷酸杂交来鉴定。空白的方框表示一个正常男性,而黑色方框表示全突变受累男性。浅灰色符号表示携带前突变的个体。符号下方或旁边的数字表示重复数目(引自 Warren ST, Nelson DL. Advances in molecular analysis of fragile X syndrome. JAMA 1994;271:536~542.)。

虑该病的遗传筛查项目。使用 FMR1 蛋白抗体进行的预先研究表明，筛查该基因产物存在与否是可能的。

脆性 X 综合征分子基础的发现和该疾病中三联体重复和动态突变的作用，不仅解释了 Sherman 悖论中特殊的遗传方式，而且随之很快地发现了其他一些遗传病的三联体动态重复(图 8-44)。强直性肌营养不良病中肌强直基因 3′非翻译区 CTG 重复的作用已在第三章中讨论过，在 Huntington 病和其他神经退行性疾病涉及的基因的编码区中编码多聚谷氨酰胺的 CAG 重复也是如此。值得注意的是，这些三核苷酸引起疾病的机制在不同疾病中明显不同。在脆性 X 综合征中，高度甲基化和转录失活导致功能丧失；在显性遗传的神经退行性疾病中，有异常的蛋白产生；而在肌强直性营养不良中，尽管发病机制尚不清楚，但它可能涉及 mRNA 稳定性的改变。就在最近，发现了一个独特的第四种三联体重复疾病。以共济失调、口吃和心脏病为特征的 Friedreich 共济失调，是一种常染色体隐性神经退行性疾病。它是由编码 Frataxin 的 *X25* 基因内含子 1 的 GAA 动态重复引起的。这种扩展在正常人为 10~20 个拷贝，在病人为 200~900 个拷贝，它可能通过产生多个 AG 剪切受体位点而干扰 Frataxin 异质核 RNA(heterogeneous nuclear RNA)，导致细胞质中成熟 mRNA 的缺乏。



图 8-44 人类四类三联体重复疾病。一个假想的基因被描述为具有 3 个外显子(淡灰色方框)和两个内含子(介于中间的淡灰色线)。转录起始位点(AUG)和终止信号(TAA)以黑色垂直栏表示。在脆性 X 综合征中，CGG 重复位于 *FMR1* 基因外显子 1 5′非翻译区。在 Friedreich 共济失调中，GAA 重复位于编码 Frataxin 的 *X25* 基因内含子 1。在 Huntington 病和所列的其他神经退行性疾病中，CAG 重复位于外显子，编码一个延长的多聚谷氨酰胺。在强直性肌营养不良中，CTG 重复位于 DM 蛋白激酶(myotonin，肌紧张素)最后一个外显子的 3′非翻译区(引自 Warren ST: The expanding world of trinucleotide repeats. Science 1996;271:1374~1375.)。

5. X 常染色体易位

相互易位偶尔也发生在 X 染色体和一条常染色体之间。这些自然事件证明在基因定位(如第九章讨论的 DMD)和理解表型-核型关系上非常有用。可能在染色体节段的断裂和重接中，X 染色体上某个基因内发生了一些染色体物质丢失或断裂。因为在 X 常染色体易位的病例中，正常 X 染色体优先失活，所以这种易位的杂合性女性可能表现 X 连锁隐性疾病。否则，这种疾病只能见于半合子的男性。

进行染色体分析的指征(表 8-5)

(1) 已知或可疑的染色体异常。带有已知或可疑染色体异常的患者通过核型分析进

行诊断或证实是必需的。例如,即使是有经验的内科医生,也可能漏诊新生儿中的 Down 综合征。这些严重疾病应当得到最精确的诊断,而且没有任何理由不进行这样的研究。正如以前讨论的那样,在 Down 综合征病例中,判断个体是 21 三体性还是从平衡易位携带者父母处继承来的非平衡携带者也是重要的。

(2) *多发的先天性异常,特别是当涉及生长和/或智力低下时*。因为严重染色体异常的特征性表型是生长迟缓、智力低下和体征异常,所以这些综合表现通常提示存在染色体异常的可能性,应当进行核型检查。轻度的异常可能需要进行前中期染色体分析或应用 FISH(针对微缺失综合征)。

(3) *性分化疾病*。我们以前在这一章讨论过的这些疾病需要进行染色体的性别判定以获得精确的诊断和处理。性染色体和常染色体的非整倍体的快速精确的诊断可以通过有丝分裂期或间期细胞的 FISH 分析来完成。

(4) *无法诊断的智力低下*。脆性 X 综合征的高发病率要求我们应该在 X 连锁的智力低下病人以及任何具有无法诊断的智力低下病人中进行该疾病的检查。除了分子诊断以外,染色体分析是判定的重要部分。

(5) *血液恶性肿瘤和与染色体不稳定性有关的疾病*。在某些血液恶性肿瘤中常见的特异性染色体异常已经确定了染色体研究在这些疾病的诊断和预后分析中的价值。另外,还有一些遗传病与染色体不稳定、白血病和其他恶性肿瘤的高发风险有关。

(6) *多发流产*。在大约 5% 有过两次或两次以上自发流产的夫妇中,一方携带平衡易位。虽然未能证实,但人们推测不平衡状态可能是这些自发流产的原因。不过,通过产前诊断对这样的夫妇进行检测是监控未来妊娠的一个指征。

表 8-5 进行核型检查的指征

已知或可疑的染色体异常
多发的先天性异常
性分化疾病
无法诊断的智力低下
选择性血液恶性肿瘤
多发流产

最后,在明确的孟德尔式遗传病中通常无需进行染色体检查,这些疾病一般没有可检测到的染色体异常。然而,在罕见病例中(见第九章),发生在单个基因内断裂的染色体易位可能产生一个显性遗传病。并且,应当记住,一个小的染色体缺失在一个家系中也可能像单个基因显性孟德尔式性状一样分离。随着更新的、提高检查染色体结构异常分辨率的技术的发展,人们发现先前认为与染色体改变无关的疾病实际上可能有这样的染色体改变。

推荐读物

Belmont JW. Genetic control of X-inactivation and processes leading to X-inactivation skewing. Am J Hum Genet 1996;58:1101 ~ 1108. *A concise, current review of the mechanism of X chromosome inactivation and its varia-*

tions.

Emanuel BS. The use of fluorescence in situ hybridization to identify human chromosomal anomalies. Growth Genet Horm 1993;9:6 ~ 12. *A clear discussion, with useful diagrams, of the rationale and uses of FISH analysis.*

Epstein CJ. Down syndrome (trisomy 21). In: Scriver CR, Beaudet AL, Sly WS. Valle D, eds. The Metabolic and Molecular Bases of Inherited Disease. 7th ed. New York: McGraw-Hill, 1995;749 ~ 794. *A thoughtful and comprehensive review of the clinical, cytogenetic, and molecular aspects of Down syndrome.*

Ferguson-Smith MA, Goodfellow PN. SRY and primary sex-reversal syndromes. In: Scriver CR, Beauder AL, Sly WS, Valle D, eds. The Metabolic and Molecular Bases of Inherited Disease. 7th ed. New York: McGraw-Hill, 1995:739 ~ 748. *A brief summary of sex reversal syndromes, and a clear explanation of the discovery of SRY.*

Gardner RJM, Sutherland GR. Chromosome Abnormalities and Genetic Counseling. New York: Oxford University Press, 1989. *A useful book for dealing with counseling issues related to chromosome abnormalities; unfortunately written before the discovery of triplet repeat disorders or of the use of FISH.*

Hassold TJ. Chromosome abnormalities human reproductive wastage. Trends Genet 1986;2:105 ~ 110. *A summary of the important role played by chromosome abnormalities in reproductive wastage.*

Jacobs PA. Human population cytogenetics: the first twenty-five years. Am J Hum Genet 1982;34:689 ~ 698. *A delightful summary of the important early years in human cytogenetics, told by an active participant and contributor.*

Le Beau MM. One FISH, two FISH, red FISH, blue FISH. Nat Genet 1996;12:341 ~ 344. *A concise discussion on the status of FISH and its applications in a variety of research and clinical areas.*

Ledbetter DH, Ballabio A. Molecular cytogenetics of contiguous gene syndromes: Mechanisms and consequences of gene dosage imbalance. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The Metabolic and Molecular Bases of Inherited Disease. 7th ed. New York: McGrawHill, 1995:811 ~ 839. *A comprehensive discussion of the molecular basis for, and diagnosis of, a number of disorders secondary to deletions and duplications resulting in alterations in gene dosage.*

Sapienza C, Hall JG. Genetic imprinting in human disease. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The Metabolic and Molecular Bases of Inherited Disease. 7th ed. New York: McGraw-Hill, 1995;437 ~ 458. *An extensive discussion of imprinting in humans and mice, and a wide-ranging discussion of the potential implications of this phenomenon in human disease.*

Tyler-Smith C, Willard HF. Mammalian chromosome structure. Curr Opin Genet Devel 1993;3:390 ~ 397. *A current review of functionally important elements of chromosome structure.*

Warren ST, Nelson DL. Advances in molecular analysis of fragile X syndrome. JAMA 1994;271:536 ~ 542. *An excellent, clear, exposition of the molecular basis of the fragile X syndrome and the current status of molecular diagnosis.*

第九章　人类基因组解剖学：基因定位、连锁及定位克隆

人们经常问我："为什么给基因定位？"原因很多，但其中三个更为突出。第一个原因与挑战有关。过去只是因为开始时有不可克服的困难激发我们中许多人去开发人类基因定位的方法。第二个动机是满足人们的审美需要。由于某种原因，把一熟知基因定位于正确的位置，我们获得了满足。然而，还有第三个：实用原因。我相信我们可以有把握地说："基因定位对你有益"！

——Frank Ruddle（引自 Reverse Genetics And Beyond. Am J Hum Genet 1984;36:944.）

正如基因定位领域早期带头人之一在上述引语中所说的，将基因定位于相应的染色体上绝大多数始于一种智力活动，而与生物学或医学无直接和明显的关系。然而，过去20年中由于基因定位技术的迅速发展使以前难以认识的疾病基因现在可以用称为定位克隆的方法克隆。

1950年以前基因定位基本上限于X染色体。那时X连锁疾病特征遗传模式通过系谱调查可以进行基因定位。然而，常染色体的基因座定位则非常困难。Donahue于1968年首次将常染色体基因Duffy血型基因座定位于染色体上。事实上，这是"功夫不负有心人"规律的极好范例。在细胞遗传学的早期Donahue正学习如何显现中期染色体，并且用他自己的白细胞试验。他注意到他的一个1号染色体有一大的异染色质区域，正好位于着丝粒的下方（图9-1）。这与任何疾病无关，并仅代表这个染色体的可变特点，通常称作"异形性"。出于好奇，他还研究了其家庭的其他成员，并观察到这个染色体变异的孟德尔式传递。作为这一分析的一部分，他还获得了该家族其他成员的血型信息。在检查这些结果时，他注意到尽管更有名的ABO和Rh血型红细胞抗原与1号染色体上的这一发现无关，但Duffy血型这种较少检测的红细胞抗原的遗传完全可由1号染色体异形性在家庭中预测出来。图9-1示系谱连同染色体发现和Duffy血型基因型。家系调查表明1号染色体异形性无一例外与Duffy基因座上的等位基因一起传递。Donahue得出正确结论：Duffy血型基因一定位于1号染色体，因而在常染色体基因定位上首次获得成功。

自从1968年Donahue的发现，分子遗传学技术的显著进展，已导致成功定位16 000多个人类基因。最近，通过大规模人类基因组计划的共同合作努力，这些基因中的绝大部分已绘入人类遗传图（见第十章）。

一、连锁分析

Donahue将Duffy血型基因座定位于1号染色体是遗传连锁分析的一个早期的例子。



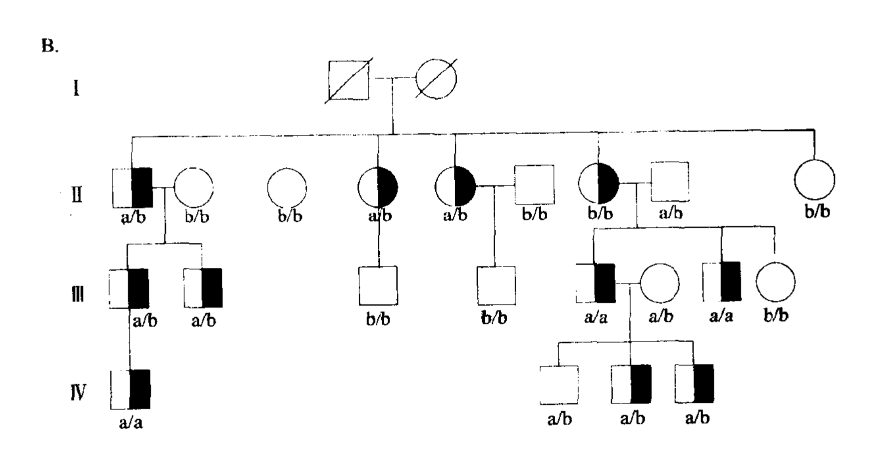


图 9-1 Duffy 血型与 1 号染色体连锁 。A 示来自两个不同个体的一对 1 号染色体。这些研究是在染色体显带出现以前所做的，在左侧一对染色体中的一条可检测到着丝粒的下方长度的明显增加，称为异形性。右侧的染色体显示更为常见的形态。B 中的系谱表示 Duffy 血型和 1 号染色体异形性的遗传模式。半实心符号表示在 1 号染色体上存在异形性，a 和 b 的标记表明 Duffy 血型的基因型。虽然第一代个体为未知状态，但从该系谱的其余部分能看到，有异形性的染色体总是与 Duffy 基因座上的等位基因一起传递(引自 Donahue RP, Bias WB, Renwick JH, McKusick VA. Probable assignment of the Duffy blood group locus to chromosome 1 in man. Proc Natl Avad Sci USA 1968;61:949～955.)。

这一方法是通过分析一个基因在相同染色体上附近的另一基因座来定位的。这样的位点就称为是“连锁的”。

Donahue 的方法很难普遍用于大量基因的定位，是由于在大多数染色体上不出现可

见的染色体异形性。现在人类疾病基因的定位主要是通过使用一系列特征性强的多态DNA标记的遗传连锁分析进行的,这比Donahue所用的单个1号染色体异形性具有更丰富的信息,然而,它们都是基于相同的基本遗传原理。

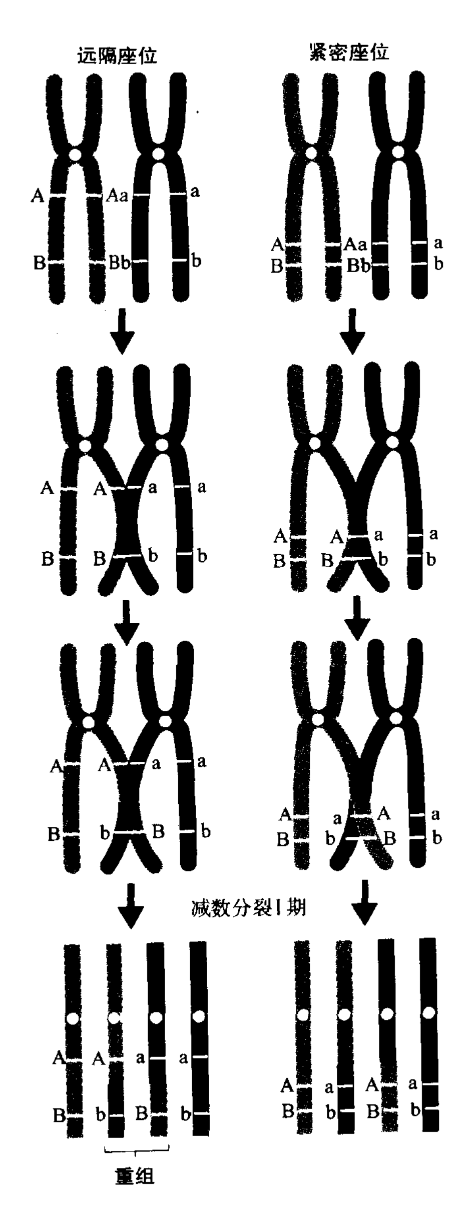


图9-2 减数分裂中同源染色体的交换。左侧示两个基因座相距较远,以致它们之间可能发生重组。右侧示两个基因座紧靠在一起以致它们之间的重组将是罕见事件。

为了理解连锁分析的概念,有必要密切掌握发生在减数分裂Ⅰ期的事件,尤其是关于交换。图9-2示一对同源染色体的减数分裂。在图的左侧部分,追踪的两个基因座指定为A(或a)和B(或b),在染色体上相距较远,这使它们之间可能发生交换。从图的底部可以看出,这种交换产生等位基因重组的染色体。在这个例子中,开始时含有等位基因A和B的染色体和含有a和b等位基因的染色体,经过减数分裂会产生共4种可能的重组。

同样的交换在图的右侧可见,此时两个基因座紧靠在一起。结果,A和B在一起,a和b在一起。正如该图所示,两个基因座彼此相距越近,它们之间重组的可能性则越小。因此。两个基因座间的重组可用来粗略地估计它们之间的距离。

图9-3示这一方法在孟德尔式疾病遗传中的应用。该方法原理简单:如果疾病基因座与一个多态基因座(通常称为"标记基因座")相距很近,而且可追踪其在家系中的遗传,就可根据该标记基因座预测谁将受累。这种分析的错误机率是标记基因座和疾病基因座分开距离远近的函数。该图中父亲患显性疾病,因此,他的基因型在疾病基因座为DN,而他的正常妻子为NN。在标记基因座,他的基因型恰好是Aa,他的妻子为AA。在他的后代中除Ⅱ-4外,所有的均显示出疾病基因座和标记基因座的完全连锁,因为继承父亲A等位基因的所有后代也继承了疾病基因,而继承a等位基因的所有后代在疾病基因座均继承了一个正常基因。实际上,人们需要这样信息来判断父亲的疾病基因是否与标记基因座的A基因位于相同的染色体上,这被称为两个基因座的"决定相"。

然而个体Ⅱ-4是个例外。该后代患此病,但是,从其父亲继承了一个a等位基因。这是两个基因座重组的例子。

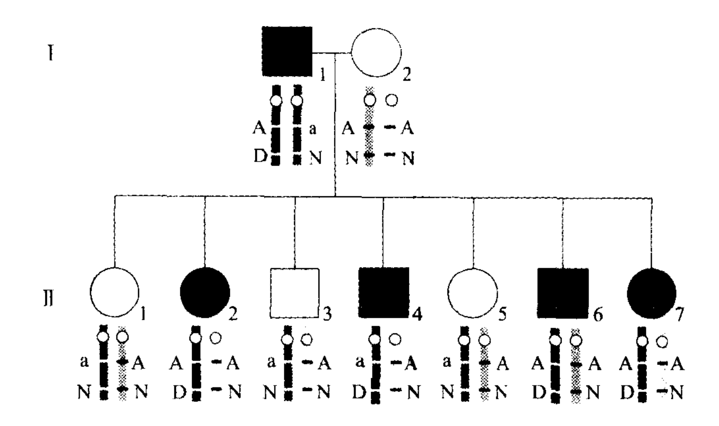


图 9-3 致病基因和标记基因座的连锁。实心标记的个体为显性遗传病患者，基因型为 DN，未受累个体为 NN。父亲（Ⅰ-1）携带致病基因的染色体上标记基因座的等位基因是 A，而他的正常染色体携带等位基因 a。在标记基因座，未受累母亲在标记基因座是纯合子 AA。下面表示 7 个后代携带的父方染色体在左侧，携带的母方染色体在右侧。个体Ⅱ-4 表明在标记基因座和致病基因间发生了重组。因其继承父亲标记基因座上的一个等位基因，而且已继承了致病基因 D。其余同胞为非重组体。

在这种分析中，重要的是要意识到并非标记基因座的等位基因本身引起此病。因此，例如在无此病既往史的其他家系中，在标记基因座发现带有 AA 基因型的个体并不表明个体患病（Ⅰ-2 就是这样的例子）。一个重要的推论是，一般只有当几个家庭成员，包括某些疾病的患者可以利用时，才可进行连锁分析以追踪致病基因的遗传。而且，如果父亲是纯合子 AA 或 aa，就不可能进行图 9-3 所示的这种分析，因为在这个例子中，人们不能用标记基因座区分父亲的两个染色体。因此，在连锁分析中，最有用的是那些高度多态的标记基因座。

确定两个基因座间距离的遗传术语是单位厘摩（cM）。如果两个基因座相距 1 厘摩，当染色体从双亲传递给孩子时，这两个基因座间的重组机率是 1%。因此，人们可以把厘摩看作染色体距离的粗略单位，虽然与物理距离的关系不呈真正的线性，因为这些区域在重组的潜能上稍有差异。每个染色体平均约为 150 个 cM；整个人类基因组大约有 3 300 个 cM，相应的碱基数是 3×10^9bp。因此，平均 1cM 大约是一百万个碱基对（1 000kb）。若是一个标记基因座与疾病基因紧密连锁，不仅一定要在相同的染色体上，而且还要位于那个染色体的同一区域。在同一染色体上距离远的标记因为两个基因座的重组率相对高，甚至可能不会出现连锁。

用于连锁分析的标记基因座可以是被研究家系中任一可变的事物，并由基因组中单一位置所决定，比如 Donahue 染色体异形性（图 9-1）。许多早期的研究使用如红细胞抗原（ABO，Rh 等）多态蛋白质。图 9-4 是一例强直性肌营养不良基因和分泌基因座的连锁关系的研究，这种关系是在 20 世纪 70 年代早期确认的。强直性肌营养不良是常染色体显性遗传病，已在第三章讨论过。分泌基因座决定血型碳水化合物结构（A、B、H）是否分泌进入唾液。图 9-4 中 A 示家系成员的分泌基因检测结果及疾病状态。分泌基因座有两个等位基因 Se 和 se，Se 等位基因是显性基因，因此“+”个体（分泌型）的基因型是 Se/Se 或

Se/se。“－”性个体（非分泌型）一定是纯合子 se/se。在图 9-4B 中，根据这个信息重建了分泌基因座的基因型。这种重建受益于这一事实，即“－”性个体的双亲一定携带 se 等位基因，这可用于家系中模棱两可成员的确定。只有个体Ⅲ-1 的基因型不能完全确定。家系调查表明携带 Se 等位基因的Ⅰ-1 个体患强直性肌营养不良，并无明确交换的例证。强直性肌营养不良基因于 1992 年被克隆，含有的三核苷酸重复序列的扩增引起此病（见第三章）。这个基因位于染色体 19q13，靠近决定分泌表型的岩藻糖基转移酶 2 基因。

多态蛋白基因位点的数量有限，因此，多数重要的疾病基因不能用这种方法定位，总的来说没有足够相邻的蛋白多态可以显示连锁关系。然而，重组 DNA 技术通过提供几乎无限 DNA 序列多态，极大地增加了遗传标记的数量，首先是由 DNA 印迹杂交最初确定的限制性片段长度多态性（RFLP），其次是由 PCR 确定的简单多态性重复序列标记（SSR）（见第五章）。现在，有 20 000 多个这样的多态标记已被鉴定并将它们定位于整个人类基因组各自相应的位置。因此，平均而言，在人类基因图上每个厘摩内有几个可用的标记。为寻找某疾病的表型与其中一个标记的连锁证据可依每 10～20cM 为间隔（160～330 个标记）形成一整套的这种标记，以对感兴趣疾病的家系进行分型。通常，这是定位克隆方

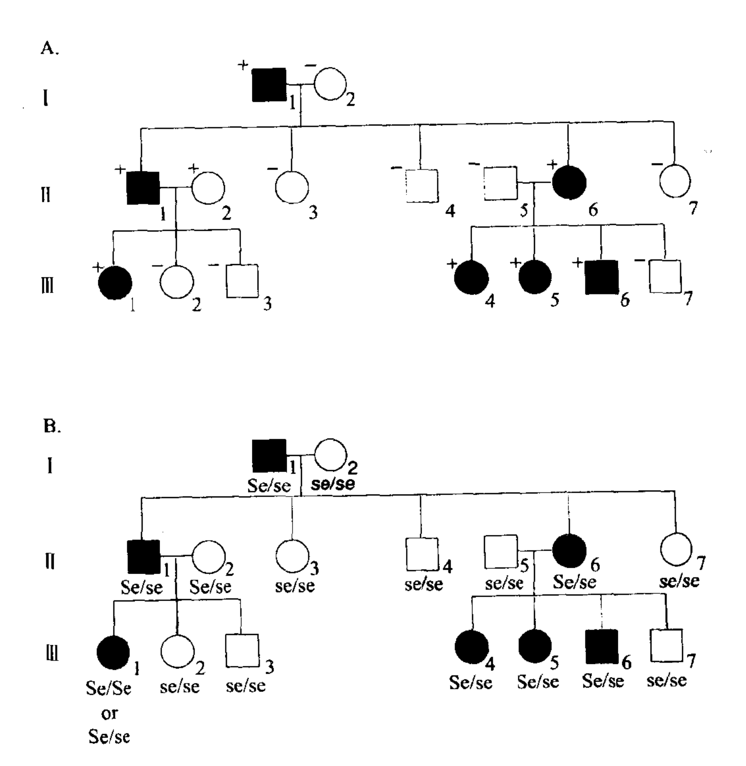


图 9-4　表示强直性肌营养不良基因与分泌基因座连锁的系谱。实心标记的个体患有常染色体显性肌营养不良症。A 示这个系谱中每一个体的分泌表型（＋或－）。基因型见 B，由于 Se 是显性的事实，因此，＋性个体基因型是 Se/Se 或 Se/se，而－性个体的基因型必是 se/se。除个体Ⅲ-1 外，所有＋性个体的基因型可以通过分泌基因座的孟德尔式遗传的必然性推断出来。例如，个体Ⅰ-1 必定是 Se/se，而不是 Se/Se，因为他有三个 se/se 基因型后代。B 表明强直性肌营养不良基因与分泌基因座的 Se 等位基因在此系谱中明显的连锁。

法的第一步，将在本章后面讨论。

为了理解连锁分析，我们有必要略为深入研究这种方法的数学方面。你也许已经想到，一个家系中可能偶然发生如上述强直性肌营养不良一样的连锁。而且，如果该标记基因座与疾病基因座连锁不太紧密，在这些被研究的家系中可能发现重组体。连锁分析可以解释这些可能性，甚至可用重组体的信息来计算遗传距离。

LOD 值

LOD 值(或 Z)是统计学名词，一般用来计算评价连锁结果的显著性。如果两个基因座确实连锁，LOD 值将对所获得实验数据的似然性与纯属偶然观察到的相同数据的似然性进行比较。似然性类似于概率，且更依赖于两个基因座间的距离。根据重组发生频率的多少(或重组分数)而测得的这个距离称为 θ。如果图 9-3 中的标记 A 与致病基因 D 的确连锁，而且与致病基因相距非常近，则能预测在标记和致病基因之间可能很少重组。在这个特殊的例子中，对应第Ⅱ代儿童的 7 个减数分裂，其中之一在标记和疾病基因间发生了重组，结果重组分数 θ 为 1/7，或约为 14 厘摩。为了计算这个似然性，必须为 θ 确定一个特定值。特定 θ 的似然性比率确定为：若两个检测基因座确实连锁而且相距距离 θ 所获得观察数据的似然性除以这两个基因座不连锁时获得相同数据的似然性。这个比率的对数(lg)被称为 LOD 值，即"优势的对数"。实际 LOD 值计算的例子见附录。实践中，LOD 值是在大范围的 θ 值基础上计算出来的。产生最高正 LOD 值的 θ 值是最可能的重组分数($\hat{\theta}$)。所报道的 LOD 值通常是最大 LOD 值($\hat{Z}$)。这就是在 $\hat{\theta}$ 处计算出来的 LOD 值。

正 LOD 值常支持连锁的存在，而负 LOD 值则表明连锁的似然性较少。通常，正 LOD 值为 +3 或更大时被认为是两个基因座连锁的确定证据；因为 LOD 值是以 10 为底的对数计算的，LOD 值 +3 表明 1 000:1 的机会，即所观察到的连锁不是偶然发生的。LOD 值 -2 (表明机会为 100:1 或更大的优势对抗连锁)时，这个 θ 值被认为是排除连锁的证据。重要的是要记住经典 LOD 值的临界值≥3 是一个随机数值，LOD 值 2.99 至 3.00 间的差异实际上并不比 3.00 至 3.01 间差异大。对于某些类型的连锁研究，尤其是用数百个标记进行复杂遗传性状分析时，这些标准也许应修改为稍高一些的临界值。

将一组非亲缘家系研究中获得的连锁数据结合起来通常是有意义的。由于在这些单独家系中连锁似然性是独立的，可用这些值的乘积联合确定似然性。记住 LOD 值是似然性比率以 10 为底的对数，而且 lg(AB)等于 lg(A) + lg(B)。因此，从单个家系获得的 LOD 值可以简单地相加以计算总的 LOD 值。在某些研究中，虽然任一单个家系 LOD 值本身不可能达到这个值，总的 LOD 值可以满足统计学显著性的标准。

虽然一个简单例子的 LOD 值可以如附录手工计算，但实际上，为此目的已开发了一组复杂的计算机程序。这些程序可以从非常复杂的家系数据中测定连续 θ 值范围的 LOD 值，同时考虑到若干其他重要因素，如疾病的不完全外显率、不同的遗传方式和潜在的基因座异质性。这些计算机分析还能结合从多个遗传标记中获得的信息，以决定它们之间最可能的排列顺序和距离。

图 9-5 描述一个常染色体显性神经纤维瘤病(*NF1*)家族的系谱,显示用一个 RFLP 标记的 DNA 印迹杂交的结果,这个标记的两个等位基因是 3.0kb 和 4.7kb。家系调查表明个体Ⅰ-1 携带的 4.7kb 等位基因随着家系中神经纤维瘤基因传递而遗传。还要注意这个 RFLP 标记本身并不检测该致病基因;个体Ⅱ-9 是 4.7kb 等位基因的纯合子,但她并不发病。连锁的一个例外是个体Ⅲ-7,他继承了父亲 3.0kb 等位基因,但却患有此病。

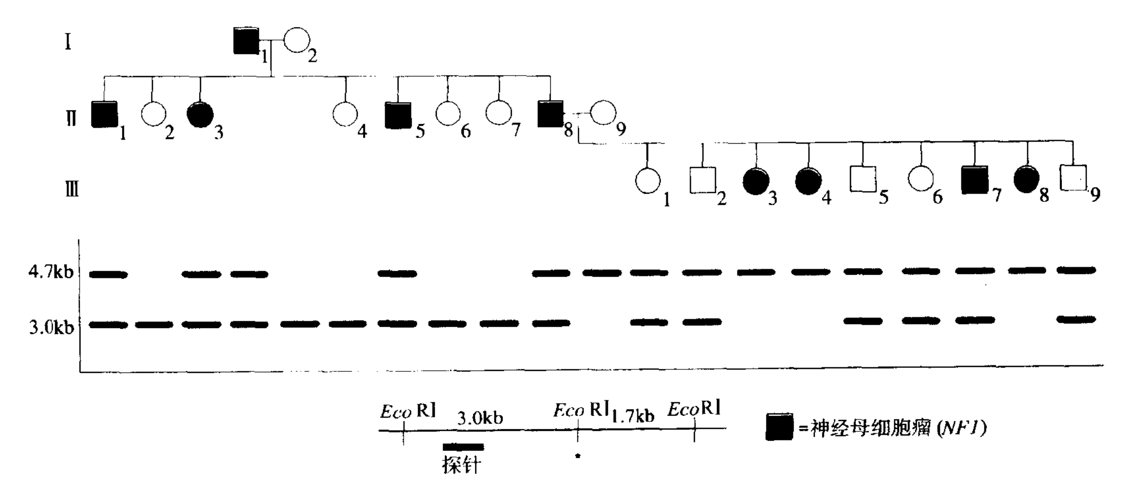
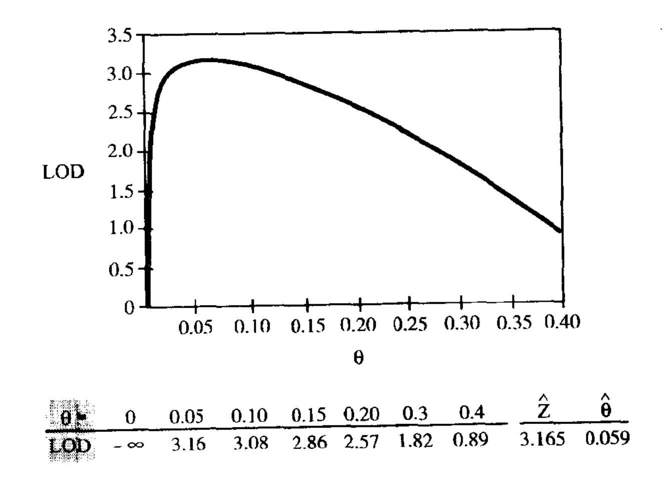


图 9-5　17 号染色体上探针 *NF1* 的连锁分析。如该图底部描述的,该探针识别带有 *Eco*RⅠ酶切位点的染色体上一个 3.0kb 的等位基因,星号表示该酶切位点。如果该位点消失,就会识别一个 4.7kb 的等位基因。系谱中每一个体的 DNA 印迹杂交结果显示在他们各自系谱符号的下面。对此标记探针而言,个体Ⅰ-1 携带的致病基因 *NF1* 是在带有 4.7kb 等位基因的染色体中传递,个体Ⅲ-7 除外,他似乎是一个重组体。



θ	0	0.05	0.10	0.15	0.20	0.3	0.4	$\hat{Z}$	$\hat{\theta}$
LOD	$-\infty$	3.16	3.08	2.86	2.57	1.82	0.89	3.165	0.059

图 9-6　在图 9-5 所示的 *NF1* 和 RFLP 标记基因座连锁分析的 LOD 值的计算机结果。用 MLINK 计算机程序计算出的 LOD 值。此坐标图表明 θ 值从 0 到 0.4 计算的 LOD 值。在 $\theta = 0.059, Z = 3.165$ 处,LOD 值为最大峰值。$\theta = 0.05$ 和 0.01 处 LOD 值的手工计算见本章附录。

虽然简单调查这个系谱无疑显示这个标记基因座与神经纤维瘤是连锁的，而且个体Ⅲ-7是重组体，计算机LOD值分析结果见图9-6。注意LOD值经许多θ值计算出来。如果$\theta = 0$，LOD值就是$-\infty$。为了理解这点，要记住LOD值是一个对数，而0的对数就是$-\infty$。因此，这个发现表明如果这个标记正好位于神经纤维瘤基因上，观察到一系列数据的似然性就是0。这种0似然性是由于在Ⅲ-7中发生了明显的重组。因为允许θ从0开始增加，所以LOD值是+值，在$\theta = 0.059, Z = 3.165$处是峰值，或约为6cM。对于更大的一些$\theta$值，LOD值开始下降，提示从这个家系的数据来看，标记基因座和神经纤维瘤病基因之间的最可能的距离为6cM。请注意图9-5系谱调查中这个距离产生一些直观的效应。最初的祖先Ⅰ-1有17个后代(包括子代和孙代)，其中只有一个显示重组。因此，重组的百分率接近$1/17 \times 100$，非常接近常规计算中得到的6%。这个数据可作为*NF1*和该标记间连锁明确的证据，因为LOD值在θ处大于3。

二、其他基因定位的方法

(一) 原位杂交

原位杂交是最直接的基因定位方法之一。该方法可通过克隆的DNA序列与显微镜载玻片上的分散的中期染色体进行分子杂交而得到视觉信号。早期研究经常用同位素标记的探针，现在最常用的是荧光标记。第八章已详细讨论过，通过应用染色体特异性探针，荧光原位杂交(FISH)可以区分出每一条人类染色体(图8-8)。如果有可用的该基因特异性基因座DNA探针，FISH方法也可以鉴定任一感兴趣的基因在染色体上的位置。用作探针的DNA片段必须有相当长度以提供足量的信号强度。通常探针DNA可来自于YAC或者其他包括所有或部分目的基因的大的基因组克隆。应用复杂的影像技术，染色体上不同DNA基因座顺序可以通过不同颜色标记的探针而测定。在一个标准的人类中期染色体涂片上，通常可以区分开1至几Mb的次序。通过研究间期核中的染色体(间期染色体DNA包装不十分紧密，更加分散)可得到50～500kb的较高分辨率。最后，一种称为纤维-FISH的人工伸展染色质纤维的技术可以分辨小到几个kb的距离。

(二) 体细胞杂交

这种基因定位方法来自在培养瓶中培养哺乳动物细胞的多年经验。**体细胞**(somatic cell)是有机体中与生殖细胞无关的任何细胞，体细胞杂种是将不同种系的细胞融合在一起。大多数用作人类基因定位的体细胞杂种可通过人与小鼠或人与仓鼠细胞融合而产生的，如图9-7所示。由于一些不完全清楚的原因，这些杂种细胞趋向丢失许多人类染色体而保留啮齿类染色体。经过几代后，细胞系通常相当稳定，并保留一套特定的人类染色体。这些人类染色体的保留是随机的。因此，就有可能产生这样杂种的嵌板，每一个克隆杂种细胞系含有不同的组合，甚至只有一条人类染色体。为使这种嵌板有用，使用者有必要知道每一个杂种细胞有哪些人类染色体，这可由细胞遗传学技术来测定。

图 9-8示一个体细胞杂种中分散的一条染色体的例子。

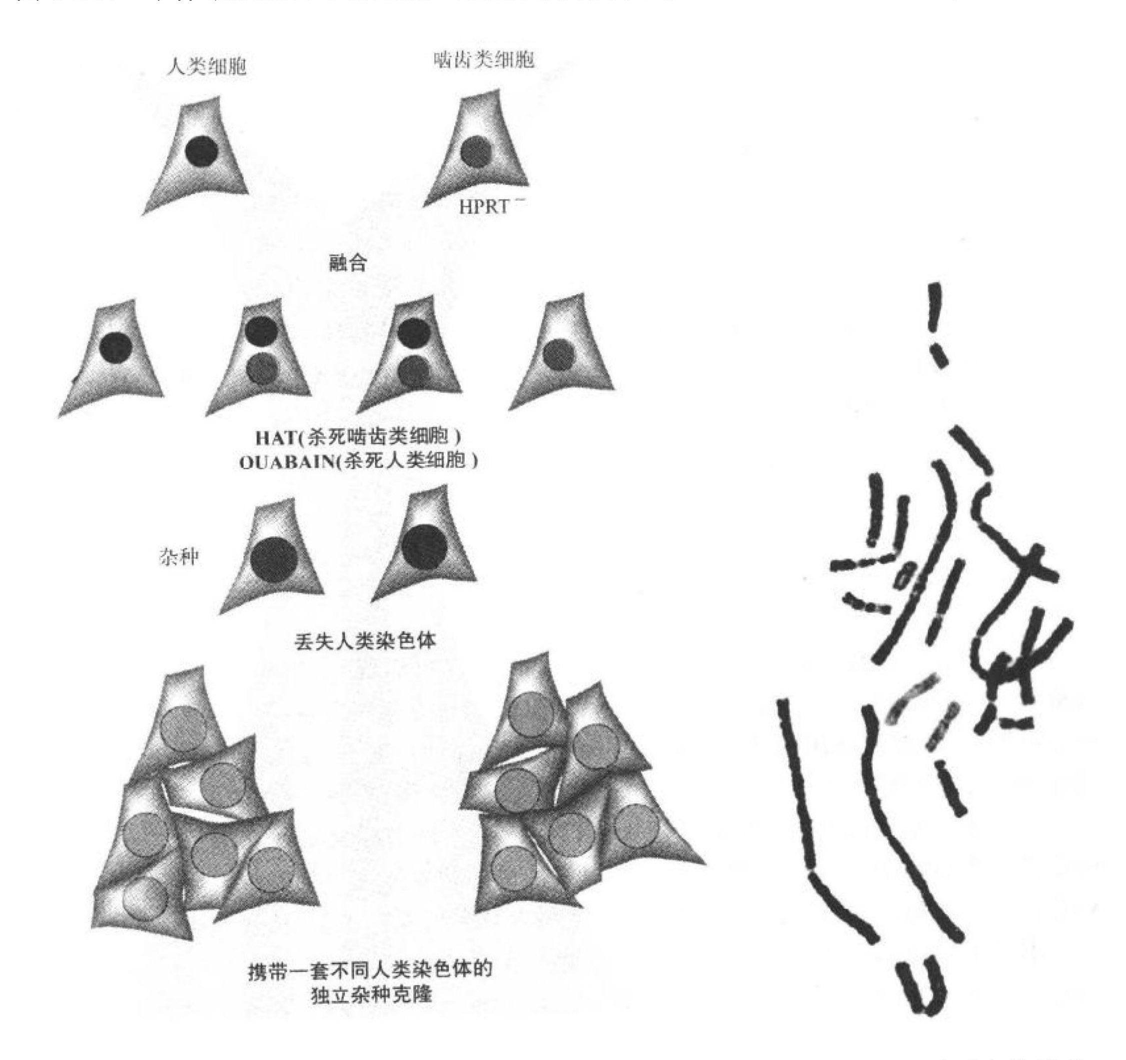


图 9-7　产生人类 - 啮齿类体细胞杂种的图解。选择含有可选择的遗传标记的啮齿类细胞,本例中啮齿类含有选择的遗传标记,次黄嘌呤磷酸核糖转移酶(HPRT)缺乏。这样可以在次黄嘌呤 - 氨基喋呤 - 胸苷(HAT)培养基中对融合后的人类和啮齿类细胞进行选择,只有具有 HPRT 基因功能的细胞可以存活。而未经融合的人类细胞也可经药物喹巴因杀死,而不能杀死啮齿类细胞。这种双重选择的结果,只有同时含有人类和啮齿类染色体的细胞才可以存活。这种杂种细胞在前几代中常优先丢失人类染色体,随后变得相对稳定。如果从大细胞群体开始操作,每个细胞克隆将包含不同的人类染色体,从而形成一个有用的体细胞杂种嵌板。

图 9-8　鼠 - 人杂种细胞核型。应用 G11 显现中期染色体,其对小鼠染色体着色比人类染色体更深一些。这个杂种的人类 3 号染色体和 X 染色体位于涂片中心附近(引自 Drs. C. Stein and T. Glover, University of Michigan.)。

这种杂种细胞嵌板一经形成,可以用于克隆 DNA 片段和表达蛋白的基因定位。这一过程的图解见图 9-9 所示。这种有意义的 DNA 片段的存在可用不同杂种细胞中的 DNA 由 DNA 印迹杂交或 PCR 进行评价,这些杂种细胞中人类染色体的互补片段以前已经测

定。通过考虑哪些染色体存在于提供阳性信号细胞系中,可以推断特异试验 DNA 片段在染色体上的位置。表 9-1 所示例中表明,肿瘤相关基因 *MET* 位于 7 号染色体上。注意表中所有含有 7 号染色体的杂种细胞为 *MET* 探针阳性,而不包含 7 号染色体的所有杂种细胞都为 *MET* 阴性。而任何其他染色体就不是如此。

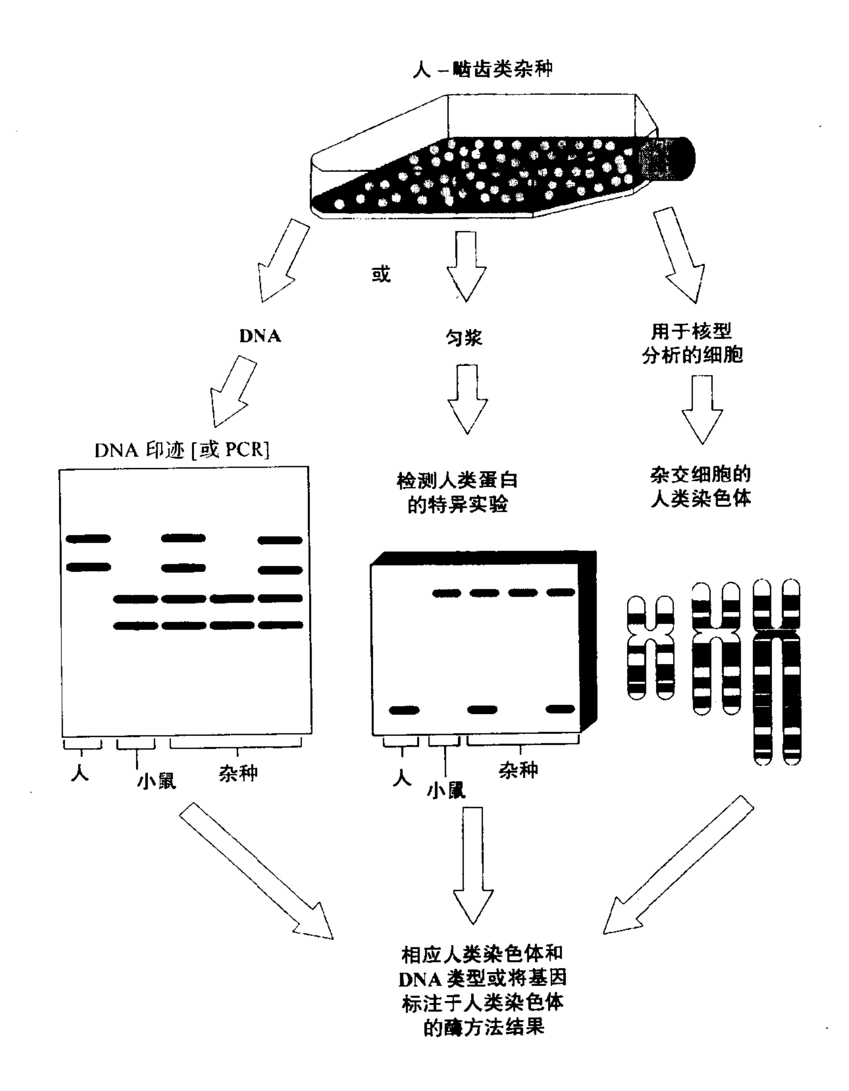


图 9-9 用人-啮齿类体细胞杂种进行基因定位。人-小鼠杂种克隆中的人类染色体内容制备如图 9-7 所示,并经核型分析确定。现在通过 DNA 印迹杂交或 PCR 识别那些含有意义序列的克隆,分析来源于个体杂种克隆的 DNA,可以进行目的基因的定位。以 DNA 印迹杂交或 PCR 反应必须识别人类序列独有的特异条带,将其同内源性小鼠基因区分开来。与此相似,只要这些蛋白在细胞中表达并且有特定的检测方法可区分人与小鼠的蛋白,也可分析来自于细胞的蛋白以确定含有人类有意义基因的杂种克隆。如本例所示,淀粉凝胶电泳实验将人类标记从内源性小鼠蛋白中区分开来。人类 DNA 序列或蛋白阳性信号的杂种细胞必然含有来源于此的特定染色体(引自 Shows TB. Sakaguchi AY. Naylir SL. Mapping the human genome, cloned genes, DNA polymorphisms, and inherited disease. Adv Hum Genet 1982;12:341 ~ 452.)。

表 9-1　采用体细胞杂交进行 DNA 探针作图

杂种	人类染色体																							人 MET
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	
A			■					■						■		■	■					■		−
B														■	■		■	■			■	■	■	−
C						■	■																	+
D		■			■			■				■							■					−
E								■																−
F	■		■		■			■	■				■	■				■		■			■	−
G	■				■			■	■					■				■		■			■	−
H	■		■		■			■	■	■				■				■		■			■	−
I	■		■		■			■	■	■			■	■				■		■			■	−
J				■			■								■		■			■				+
K				■			■				■	■				■				■		■		+
L														■										−

体细胞杂种嵌板(A-L),在啮齿类背景中每种含 1-11 条人类染色体,已经用于评估克隆的人类 *MET* 癌基因的存在。在杂种 C、J 和 K 中,*MET* 探针阳性,这仅当 *MET* 探针在 7 号染色体上时发生(经 Cooper CS, Park M, Blair DG 等允许引自。Molecular cloning of a new transforming gene from a chemically transformed human cell line. Nature 1984; 311: 29 ~ 33. © 1984, Macmillan Magazines Ltd.)。

如图 9-9 所示,只要某种可识别的蛋白实际上在杂种细胞自身中表达,杂种定位嵌板也可用于这种蛋白的基因定位。这也适用于涉及多种"看家"功能的蛋白,例如许多特殊的参与中间代谢酶类基因。对 DNA 和蛋白定位,重要的是要确定能区分啮齿类和人类的杂交信号。

在一条染色体上进行区域定位时,只要是特定染色体片段而非整条染色体出现于杂种细胞,即可应用体细胞杂种。如果有必要进行特定的区域定位,通常可以利用易位染色体(见第八章)的优点构建这样的杂种细胞。

通过生产体细胞杂种系嵌板可获得进一步精细基因定位的方法。这种杂种系含有很小的、靠辐射产生的人类 DNA 亚染色体片段(见图 9-10)。这种片段断裂的程度可由 X 线的剂量控制。与啮齿类细胞融合后,一套有限的人类染色体片段将会整合到啮齿类染色体中,随后在相应细胞系中稳定携带。这种杂种细胞称为辐射杂种。用 Southern 印迹杂交或 PCR 来筛查大型辐射杂种细胞嵌板以确定两个感兴趣基因的存在见图 9-10 所示,这两个基因在同一细胞系中出现的频率将取决于它们原来在人类染色体间距离有多近。这两个基因相距越远,则被辐射切割方法分离的可能性越大。与此相反,彼此相距很近的基因几乎总是位于同一个片段中。这种分析被认为是特殊形式的连锁。既然每个杂种细胞系含有来自于多种染色体的许多不同的人类染色体片段,解释这些数据是很复杂的并且需要繁琐的计算机程序。然而,这种方法可提供大量基因相互位置的强有力的信息,已成为人类基因组计划的重要手段。

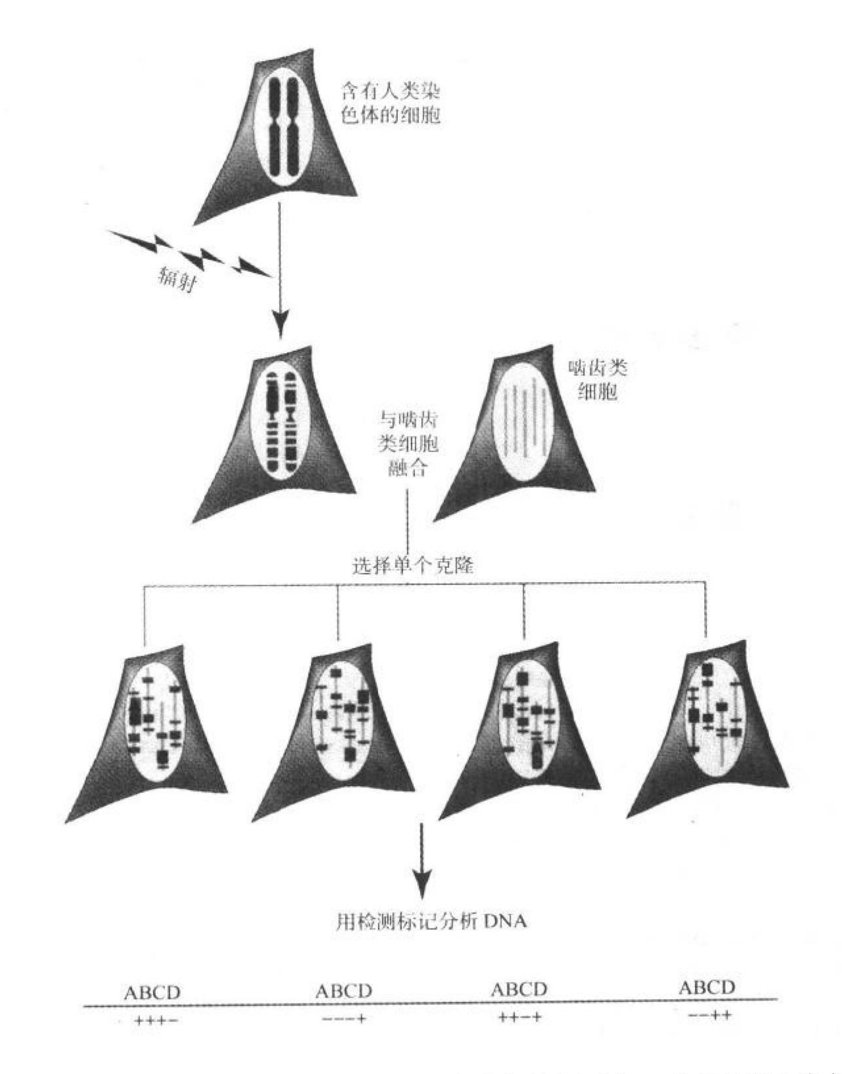


图 9-10 辐射杂种定位嵌板的制备。辐射性杂种的制备与图 9-7 描述的整个染色体体细胞杂种过程相似。可以全套人类染色体为来源，或者从仅含有一条或几条人类染色体的体细胞杂种着手。人类染色体由放射线切开，较大剂量辐射可产生较小片段。在与啮齿细胞融合后多数人类染色体片段可以稳定地整合入啮齿类染色体。每一种由此产生的体细胞杂种携带许多来自多条染色体的人类片段。两个在人类染色体上相距很近的基因更可能位于同一片段内。两个紧密相邻的基因座，用“A”和“B”表示。几条总是出现于同一个片段中，而相距较远的位点，如“C”和“D”在辐射过程中通常被随机切割分离。因此与连锁分析非常相似，两个遗传基因座出现于同一个体细胞杂种克隆中的频率与它们在基因组中物理距离成反比关系。实际上，一个辐射杂种定位嵌板通常包括几百个单个克隆。

（三）剂量效应法

如果能得到带有可变数目的特定染色体的细胞系，也有可能将某一基因定位于某一特定的染色体上。这种方法对于 X 染色体最为实用，可利用来自于含有一条至四条 X 染

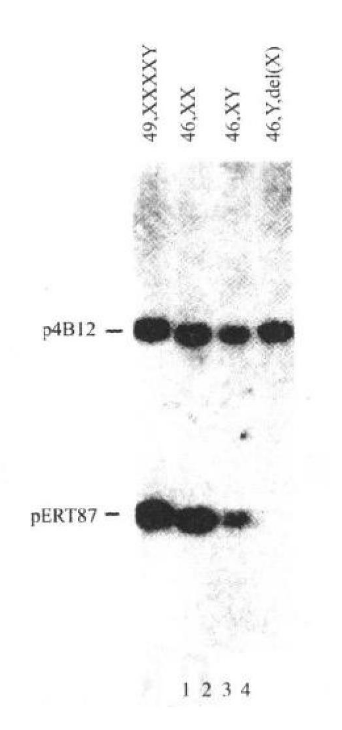


图 9-11 应用剂量效应法将一个 DNA 探针定位于 X 染色体。泳道 1 中的 DNA 来自核型为 49,XXXXY 的个体,而泳道 2 来自正常女性个体,泳道 3 来自正常男性个体,泳道 4 来自 X 染色体短臂有小缺失(见图 9.19)的男性个体(B.B.)。p4B12 探针位于常染色体上,因此在所有四个泳道中检测到相似强度的信号。pERT87 探针位于 X 染色体短臂且在 B.B.的缺失部位。可以看出在其他三条泳道中的信号强度是 X 染色体的数目函数,由此证实该探针位于 X 染色体上(经 Kunkel LM, Monaco AP. Middlesworth W 等允许引自 Specific cloning of DNA fragment absent from the DNA of a male patient with an X chromosome deletion. Proc Natl Acad Sci USA 1985; 82: 4778 ~ 4782.)。

色体个体的细胞系,如图 9-11 所示。如果使 DNA 杂交中每条泳道的 DNA 的量精确地保持恒定,位于 X 染色体上的探针(pERT87)在 49,XXXXY 细胞系中的强度将是 46,XY 细胞系的 4 倍。另一方面一种常染色体探针(p4B12)将在所有这些泳道中看上去相同。

三、人类主要组织相容性复合物

我们已经考虑了人类多态性基因座的一般特征及其在基因定位中的用途。最初在第四章介绍的人类主要组织相容性复合物(MHC,又常称为 HLA,即人类白细胞抗原),由于其高度多态性及与疾病,尤其是免疫性疾病相关,应予特殊关注。图 9-12 以图表形式解释了 *HLA* 基因座的结构。人类 MHC 位于 6 号染色体短臂,约占 4 Mb DNA。

MHC Ⅰ类和Ⅱ类基因编码细胞表面密切相关的蛋白,属免疫球蛋白超基因家族。它们作为免疫系统的抗原,分别参与自身和外源蛋白的加工和递呈过程。这些功能对于 T 细胞免疫应答调控及 B 细胞产生抗体是必需的。*HLA* Ⅲ类区域位于Ⅰ类和Ⅱ类分子间,约 1Mb,至少含有 12 个基因,其中许多基因在宿主排斥反应过程中发挥作用,包括 3 个补体基因,以及编码细胞因子肿瘤坏死因子 α 和 β 的基因,至少有 23 个Ⅱ类基因和 15 个Ⅰ类基因,但其中许多基因的功能尚不清楚。

Ⅰ类和Ⅱ类基因是哺乳类基因组中最具有多态性的基因座。至少有 57 个 HLA-A 变异体,111 个 HLA-B 变异体和 34 个 HLA-C 变异体。HLA 基因的高度多态性是通过自然选择来维持的,是不同的等位基因对于传染病病原菌抵抗力存在差异的结果。在怀孕期间,母亲和孩子的 *HLA* 基因的差异也是维持胎儿和胎盘组织的一个重要因素。

HLA 等位基因的遗传精确地遵循共显性孟德尔系统。图 9-13 为一个典型遗传方式的核心家系。*HLA* 分型应用于一些临床事件,在器官移植确定候选供者时至关重要。在这种分析中,常规地确定 *HLA-A*、*HLA-B*、*HLA-C* 和 *HLA-DR*(Ⅱ类)基因座的特定等位基因。最初的 *HLA* 分析主要是应用鉴定每个等位基因的特异抗体进行血清分型。这些抗体主要来源于多次怀孕的妇女血清,其中已产生了针对胎儿外源 HLA 抗原的抗体。血清

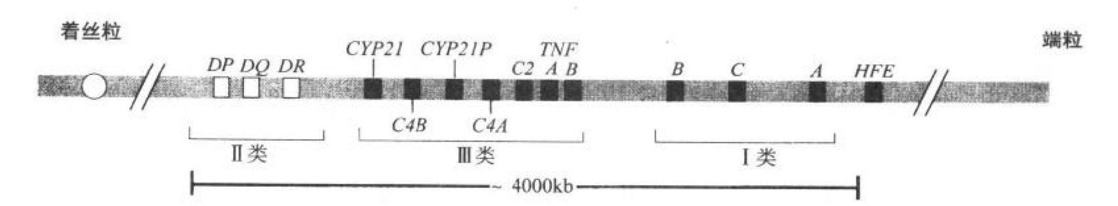


图 9-12 位于 6 号染色体短臂的人类主要组织相容复合物基因图解。尽管未按比例绘制，这里说明了该复合物的一些重要基因的相对位置，包括 HLA-A、B、C、Ⅰ类基因和 *DP*、*DQ*、*DR* Ⅱ类基因。人们早已了解的引起先天性肾上腺增生的 21 - 羟化酶缺乏与 HLA 连锁。该病的分子基础常由于位于Ⅲ类区域 21 羟化酶基因 *CYP21* 与其附近高度同源的假基因 *CYP21P* 间重排。*C2*、*C4A*、*C4B* 是补体系统成分、*TNFA* 和 *B* 基因编码细胞因子肿瘤坏死因子 α 和 β(后者现在称为淋巴细胞毒素 A)，*HFE* 是遗传性血色素沉着病基因。

分型技术逐渐被以 DNA 为基础的分析技术取代。由于最靠近端粒(HLA-A)和最靠近着丝粒(DR)间的遗传距离约为 3cM(和已知约为 3Mb 的物理距离非常相符)，因此，它们之间经常发生交换。这样的重组事件估计约占减数分裂的 3%。

如图 9-13 所示，通过调查一个家系内的 *HLA* 分型，通常有可能推断出每个 *HLA* 基因的等位基因来源于父方或母方，并用此方法判断在亲本染色体上它们如何联系在一起。

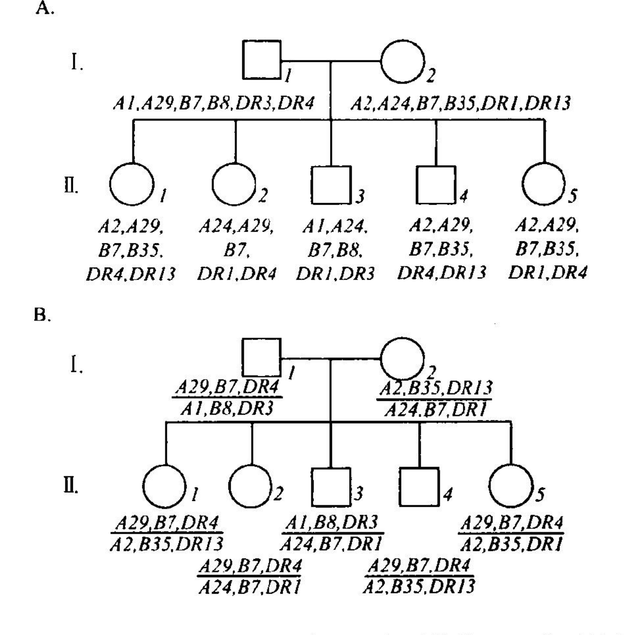


图 9-13 *HLA* 分型数据分析。A 示两个双亲和五个后代的 *HLA* 分型结果。因为人们必须判断 *A*、*B* 和 *DR* 等位基因哪些位于同一条染色体上，而分型本身不能完全确定基因型。这称为“单体型”，已在 B 中完成。通过检查后代和双亲存在的等位基因而完成单体型分析。例如，个体Ⅱ-1 一定从她的父亲继承了 *A*29、*B*7 和 *DR*4 等位基因，而从她的母亲继承的 *A*2、*B*35 和 *DR*13 等位基因，这就证明 *A*29，*B*7，*DR*4 是个体Ⅰ-1 的单体型之一；结果 *A*1、*B*8 和 *DR*3 是另一个单体型。根据相似的原因可以毫不含糊地确定这个家系中所有个体的单体型。Ⅱ-5 说明母亲单体型在 *B*35 和 *DR*13 间发生了重组，这个图仅表示了 *A*、*B* 和 *DR* 基因座的 *HLA* 分型，但在临床实践中也常可获得 *C* 和 *DQ* 基因座的分型信息。

例如,由于Ⅱ-1个体从她父亲那里继承 *A*29、*B*7 和 *DR*4,Ⅱ-3个体从父亲那里继承了 *A*1、*B*8 和 *DR*3,我们可以推断出父亲的一条染色体携带 *A*29、*B*7 和 *DR*4,而另一条携带 *A*1、*B*8 和 *DR*3。这样一套共同遗传的多态性连锁的等位基因称为**单体型**(haplotype),可以很容易地推断出该家系的4种亲本单体型。请注意,除了一个特例外,每个单体型均完整地通过减数分裂传递给第Ⅱ代。个体Ⅱ-5从其母亲那里继承了 *A*2、*B*35 和 *DR*1 等位基因,可能是一个重组体。该事件交换了两条母亲染色体上的物质,且一定发生在 *B* 和 *DR* 基因之间的某处。

Ⅱ-1个体正患急性白血病,急需骨髓移植。患者和骨髓供体在 *HLA* 基因座的完全匹配是决定移植成功的一个重要因素。如图 9-13*B*,*HLA* 单体型的数据分析表明患者的兄弟Ⅱ-4是完全匹配,他与患者从父母双方继承了完全一致的 *HLA* 单体型。Ⅱ-4称为 *HLA* 一致性同胞匹配者。个体Ⅱ-5也继承了两个相同的单体型,但由于母亲等位基因的重组事件而非最佳的匹配者。同胞Ⅱ-2有相同的父方单体型而无相同母方的,称为**单体型同一**(haploidendical)。同胞Ⅱ-3从双亲继承相反的等位基因,因而是一个完全错配者。任何两个同胞 *HLA* 完全一致的概率为25%。尽管单体型同一的同胞在某些时候也可作为骨髓供体,完全匹配仍是最佳选择。尽管不太关键,*HLA* 配型对肾脏和其他器官移植也很重要。

因为 *HLA* 复合物是高度多态的,它成为连锁研究的理想标记。而且如前所述,各种等位基因可以通过一组开展于30年前,在相关的基因尚未被克隆前的血清抗体来检测。在过去几十年中已应用 *HLA* 作为标记进行许多遗传分析,并检测一系列疾病。正如所料,大多数疾病未表现出连锁。然而,有几个重要的例子,包括下面描述的血色素沉着症,与 *HLA* 连锁是发现相应致病基因过程中的重要步骤。

血色素沉着症是一种常染色体隐性疾病,成人患者表现为过多的铁离子沉积。纯合个体在他们生命中以高速率从消化道吸收铁离子。机体铁离子的正常调控可在消化道的吸收水平进行,但正常情况下铁离子储备充足时减缓吸收的作用机制却发生缺陷。过剩的铁离子可以在心脏、肝脏、胰腺、皮肤及分泌系统沉积,导致出现各种症状,包括心脏衰竭、肝硬化、垂体激素缺陷及"血色素沉着症"(皮肤色素沉着,胰腺内分泌缺陷)。女性患者直到绝经后才出现症状,因经血流失可以防止铁离子沉积。然而男性通常在中年就出现明显的症状。最有效的治疗方法是定期移走外周血,因为它是除去铁离子的一条简单途径。因为这种方法常能成功,故漏诊通常是一个悲剧。因为疾病是隐性的,当一个个体被诊断后,检查风险家庭中的其他成员,尤其同胞是必要的。

如图 9-12 所示,特发性血色素沉着症基因(*HFE*)位于 MHC 附近,即Ⅰ类分子 A 基因的末端。甚至在血色素沉着症基因本身被克隆以前,常可通过与 *HLA* 连锁评估风险。在一个已知的有多名患者的家系内,应用与图 9-5 所示的 *NF1* 和多态标记相同的原理,可以通过特定的 *HLA* 单体型的遗传方式追踪血色素沉着症基因的分离。由于 *HLA* 基因座的高度多态性,一般可以在任何一个有一个或更多患者的家系中进行信息分析。此外,早在1975年就已发现特定的 *HLA* 类型和血色素沉着症间存在很强的相关性。而且,大多数血色素沉着症突变等位基因通常位于带有 A3 *HLA* 等位基因的染色体上。致病基因与连锁标记的特定等位基因的优先联合称为**连锁不平衡**(linkage disequilibrium)。这种连

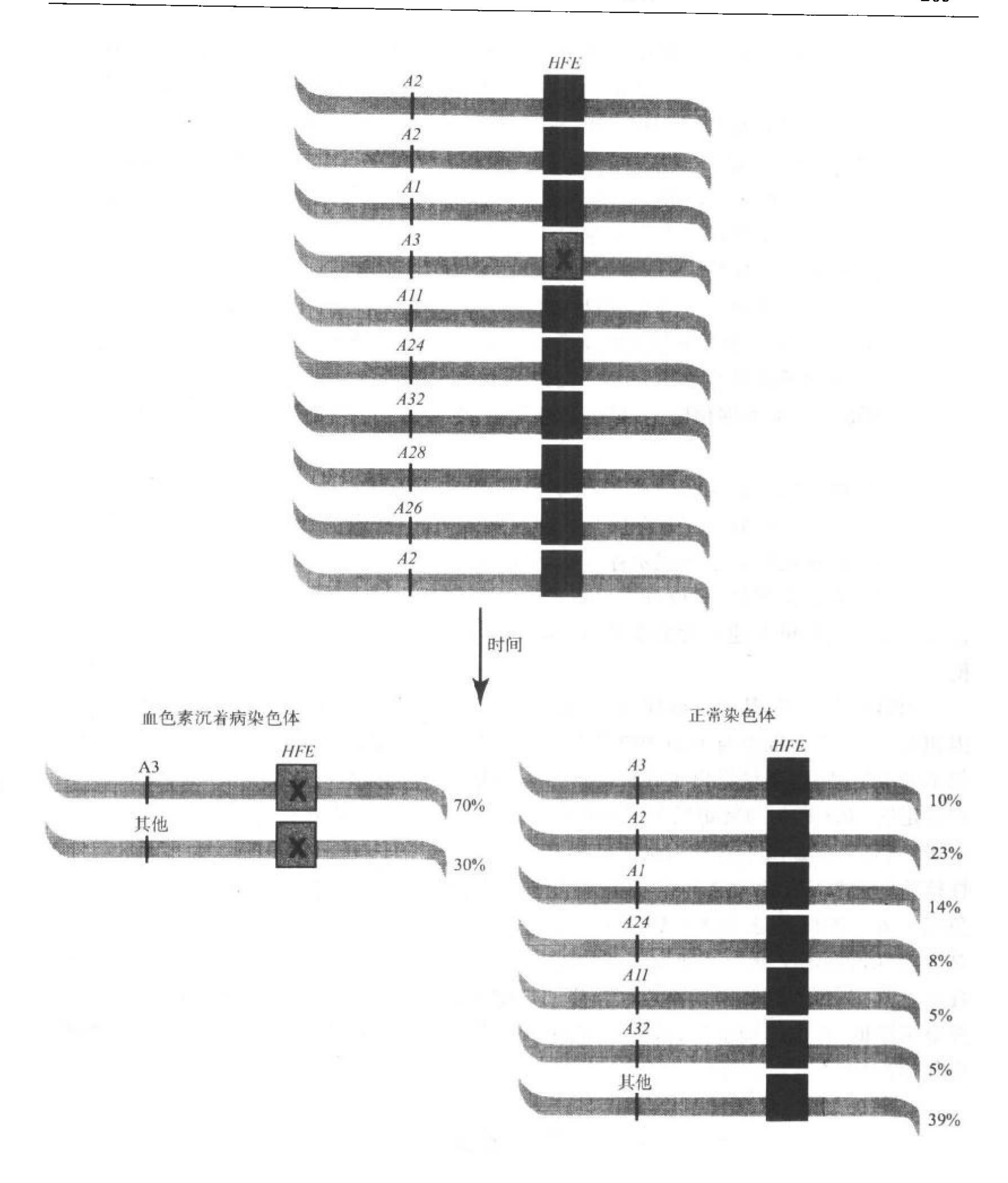


图 9-14 连锁不平衡的发生机制。在这个例子中血色素沉着症基因 *HFE* 突变(以 X 表示)仅在携带 HLA *A3* 等位基因的祖先染色体上发生一次。突变发生后的几代中,所有带有突变的染色体也带有 *A3* 等位基因。随着时间的推移,在这两个基因座间很少发生重组,以至于尽管 HLA *A3* 等位基因仅占正常染色体的 10%,但是在 70%血色素沉着症染色体上,*HFE* 基因突变仍与 HLA *A3* 等位基因关联。连锁不平衡程度与两个基因座间的遗传距离以及最初突变发生后传递的次数有关。可以预测,其他一些与 HFE 基因相距较远的多态标记,如 *HLA B*、*C*、*DR* 基因座与血色素沉着症显示出的连锁不平衡的证据非常弱。

锁不平衡可能表明大多数带有血色素沉着症突变的染色体是从一个共同祖先传递下来的。发生连锁不平衡的机制在图 9-14 中以图解说明。这种不平衡通常提示(a)大多数疾病染色体携带相同的突变。(b)检测的标记与致病基因很近。在整个 HLA 基因座存在相当多的连锁不平衡。*A3* 等位基因通常与 *B7* 和 *B14* 存在连锁不平衡，结果 B7 和 B14 也与血色素沉着症高度相关。因此，即使其他家系成员无法进行正常的连锁分析，单独 *HLA* 分型就能显著改变血色素沉着症风险的估计。

最近血色素沉着症基因(称 *HFE*)已被克隆。正如连锁不平衡检验所预测的，85%血色素沉着症染色体含有一个突变(282 位密码子的半胱氨酸替换为酪氨酸，称为 C282Y)。该突变的 DNA 分析提供了直接检测这种极其常见病(约 1:500)的方法。尽管有可信的检测方法提示铁离子过量存在，但 DNA 检验可以在铁离子过量出现以前检测到疾病的存在，这就保证了这种能够预防的、可能治愈的疾病得到理想的处理。因此，广泛的群体筛查是值得考虑的。

由于血色素沉着症纯合子($q2$)频率为 1:500，预期在群体中杂合子($2pq$)的频率为 8.5%。除非患者带有另一种并存形式的肝脏疾病如肝炎，该突变的杂合状态看来好像不伴有明显的临床症状。已有提示存在由于铁离子储备轻微升高的血色素沉着症杂合子，过去可能已有选择优势，这或许可以解释这种等位基因在群体中频率高的原因。正性选择压力可能也在单个建立者突变等位基因迅速扩散中起作用，导致出现上述的连锁不平衡。

如第四章所述，其他一些疾病，尤其是几种“自身免疫性”疾病，与特定的 *HLA* 等位基因相关，强直性脊椎炎与 HLA-B27 等位基因就是一个明显的例子。尽管这些关联的意义仍不完全清楚，但是已提出了几种假说。一种是 HLA 蛋白本身与疾病的易感性有关。如果特定的 *HLA* 产物与特定的外源性因子，如病毒重组，不能诱导有效的免疫应答，导致发生慢性感染，就会发生这种情况。如果在有某种特定 *HLA* 等位基因的环境背景中对外源性抗原的免疫应答和宿主以某种方式交叉反应，并导致正常细胞的破坏，就会产生自动免疫。另一种机制是 *HLA* 基因座本身与疾病的易感性无关，但这种现象是由于该簇内与 *HLA* 存在连锁不平衡的其他未知基因造成的。后一种假说提示自身免疫性疾病与自发性血色素沉着症相似，除了不遵循严格的孟德尔式遗传方式。最低限度人们不得不设想外显率降低，其他修饰基因，或者与环境相互作用情况的存在。现在不可能在各种假说中作出判别。

四、鉴定人类致病基因

直到 20 世纪 80 年代中期，研究人类遗传病的科学家只能通过仔细分析患者疾病的病理生理学期望发现易感基因。血红蛋白基因(第六章已详述)突变引起血红蛋白病是第一个在分子遗传学水平定性的人类疾病。血红蛋白病特别便于研究，因为疾病的病理学可以定位于红细胞，相应的血红蛋白可以通过收集外周血无限量地得到。直到最近，研究人类遗传病的标准范例都是遵循这种方式，并开始于检测患者功能缺陷。描述缺陷蛋白或者酶类的特征可随后用第五章概括的 cDNA 或基因组克隆方法克隆易感基因。这

种常用方法即从基本生化缺陷转到易感基因,已称为"功能克隆"(图 9-15)。不幸的是,这种方法仅能应用于有限的、可以直接测定生化缺陷的人类致病基因。到 1986 年,仅测出约 200 种人类疾病的特异致病基因,许多最重要的基因仍未得到解决。到 20 世纪 80 年代后期,现代分子遗传学强有力的技术手段使仅通过图谱位置定位人类疾病易感基因成为可能,而不必借助任何功能信息。这种方法称为"定位克隆"(图 9-15)。

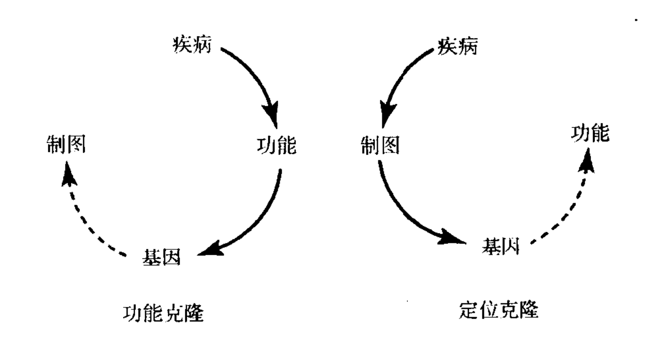


图 9-15 功能克隆和定位克隆步骤的比较。在传统的功能克隆中,基因功能的研究先于基因鉴定。在定位克隆中,基因定位在先,最后再鉴定基因。基因已鉴定后,可以进行基因功能的选择性研究。实心箭头为方法中必需步骤,点状箭头为可选择步骤(引自 Collins FS. Positional cloning: Let's not call it reverse anymore. Nat Genet 1992,1:3~6.)。

五、定位克隆

定位克隆问题的重要性见图 9-16 说明。人类基因组共约为 3×10^9bp,估计约含 100 000个基因。一条染色体含 $50 \times 10^6 \sim 260 \times 10^6$bp,相当约 1 500~8 000 个基因。定位克隆开始于整个基因组水平(3×10^9bp),止于单个碱基对,可以发现引起人类疾病的微小突变。这个过程有两个主要瓶颈。第一是应用细胞遗传学或连锁分析方法将基因定位于特定的染色体区域。一般定位于一个含几百万碱基对的区域内,通常含近 100 个基因。第二个主要障碍是从该区域内所有可能候选基因中找出正确的基因。

定位克隆鉴定的第一批基因利用了特定基因座的染色体畸变。Duchenne 肌营养不良(DMD)基因的克隆是一个重要的例子。

(一) Duchenne 肌营养不良(DMD)

Duchenne 肌营养不良(DMD)为 X 连锁疾病,多见于儿童早期的男孩,初期症状为疲乏无力。这些男孩多在 10 岁前常坐轮椅生活,大约 20 岁时死于肌肉复性的退行性并发症。

由于 DMD 是 X 连锁隐性疾病,女性患者罕见。可能的解释包括致病基因纯合子或者 X 染色体失活(见第八章)。第三种可能是通过几个女性患者细胞遗传学研究发现 X 常染色体易位,每个病例受累的常染色体不同,但在 X 染色体上的断裂点总是 p21 区域(图 9-17)。对这种现象最简单的解释是这些病例的易位断裂实际发生在 DMD 基因内。因此,人们不禁怀疑,既然这些女患还有一条正常 X 染色体,为什么她们会发病?原因在于 X 染色体失活

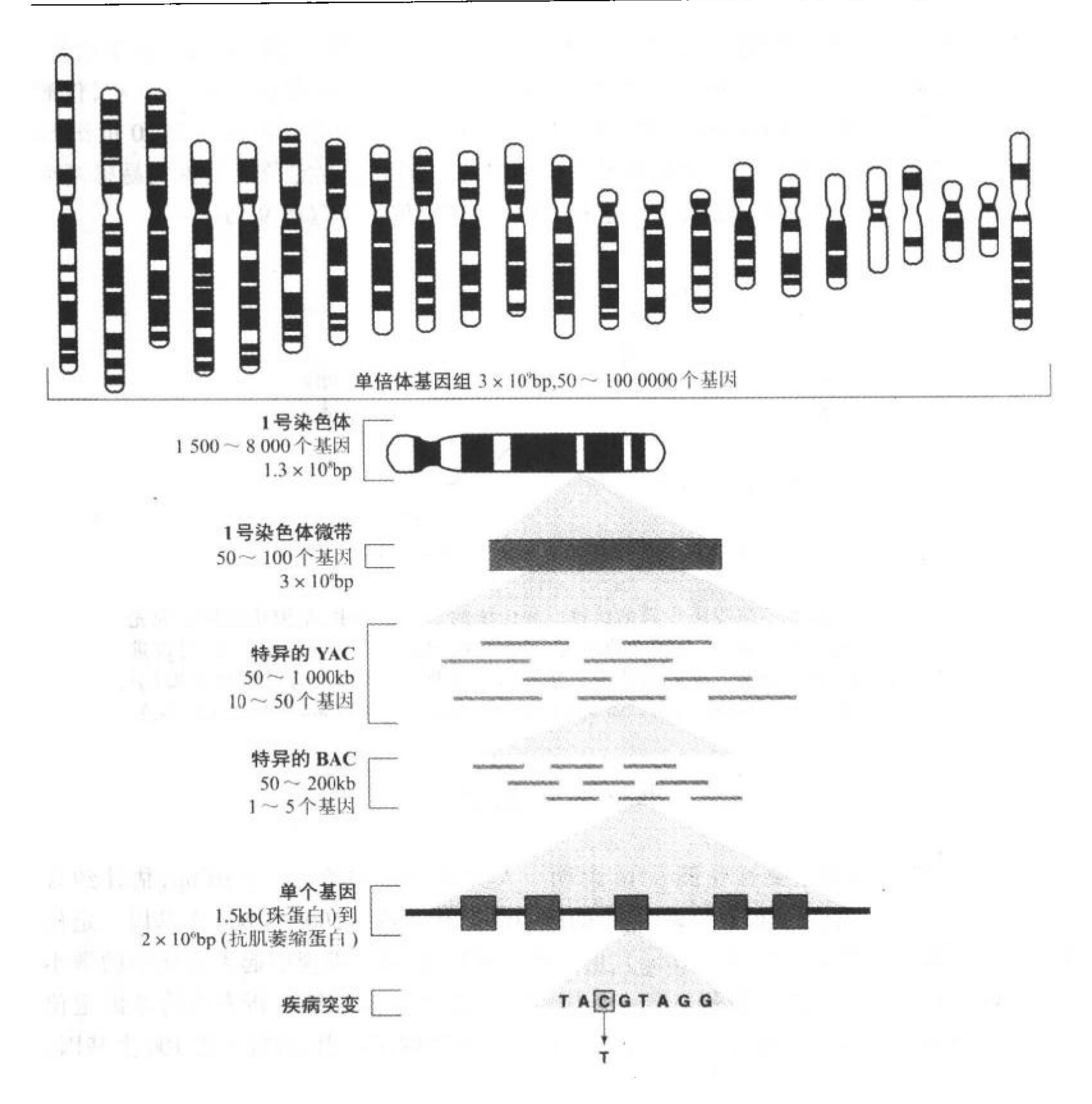


图 9-16　定位克隆的等级问题。定位克隆开始于如图中上部所描绘的人类基因组单倍体的 23 条染色体，意图深入至如图下部的单核苷酸解析，这是鉴定致病基因突变所必需的。**酵母人工染色体**(YAC)是在酵母中生长的专用克隆载体，可以携带 50～1 000kb 外源 DNA。YAC 对于大区域 DNA 定位很有价值。细菌人工染色体(BAC)是可以携带 50～200kb DNA 并在细菌中生长的专用克隆载体。尽管 BAC 比 YAC 小，但是 BAC 比 YAC 有许多技术上的优点，因而它是定位克隆又一重要工具。

现象，在所有病例中，正常的 X 染色体失活而易位的 X 染色体具有活性，导致在 DMD 位点处的正常基因为无功能性拷贝(图 9-18)。因此，这些罕见个体的发现促使人们推断出 DMD 基因最有可能定位于 Xp21，若干其他家系的经典连锁分析结果也支持这一结论。

不仅易位，而且缺失也有助于人类致病基因的定位。DMD 再次提供一个有教育意义的例子。即使是很小的染色体缺失也可能涉及一些基因(见第八章，邻近基因综合征)，所以一个染色体缺失出现的可能性，常是一个人出现多种遗传缺陷的线索。一个无任何遗传缺陷家族史的男孩(B.B.)被发现患有 DMD、慢性肉芽肿(CGD，为一种白细胞缺陷，

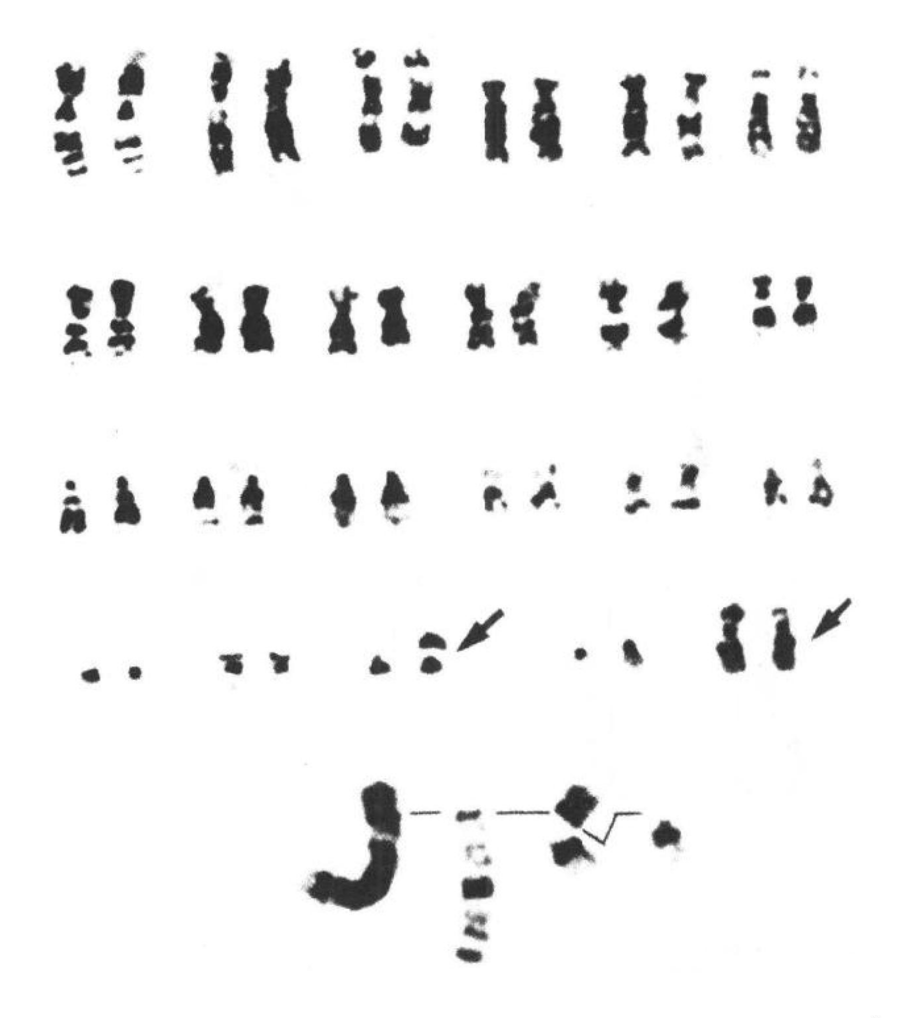


图 9-17 DMD 女性患者的核型。如箭头所示，该女性的 21 号染色体和一条 X 染色体存在明显平衡易位。底部箭头所示为高分辨图片。左侧为正常 X 染色体，邻近的为衍生的 X 染色体，其短臂的上半部与 21 号染色体的一部分发生了交换。接下来的是另一个对应的易位，其 21 号染色体短臂与 Xp 尾部相连。右侧为正常的 21 号染色体。X 染色体上的断裂点为 p21，结果证明是 DMD 基因所在的位置（经 Verellen-Dumoulin C, Freund M, DeMeyer R 等允许引自 Expression of an X-linked muscular dystroplly in a female due to translocation involving Xp21 and non-random inactivation of the normal X chromosome. Hum Genet 1984;67:115 ~ 119.）。

严重影响抗感染能力）和其他两种 X 连锁性状。这四种 X 连锁疾病的同时发生引起仔细的细胞遗传学研究，最终发现了 Xp21.2（图 9-19）的微小缺失。因为双亲均未携带这种缺失，这似乎是真的发生，使得这四种 X 连锁疾病定位于 X 染色体的相应区域。这个独一无二的个体同时也为 *CGD* 和 *DMD* 基因克隆过程中的难题提供了关键性的线索，并开始了对这两种疾病分子水平的理解。

两个研究小组分别采用两种不同的方法克隆了 *DMD* 基因。一组通过分析图 9-17 中 X 常染色体易位，克隆了该基因的一部分，该易位将 X 染色体序列和 21 号染色体短臂

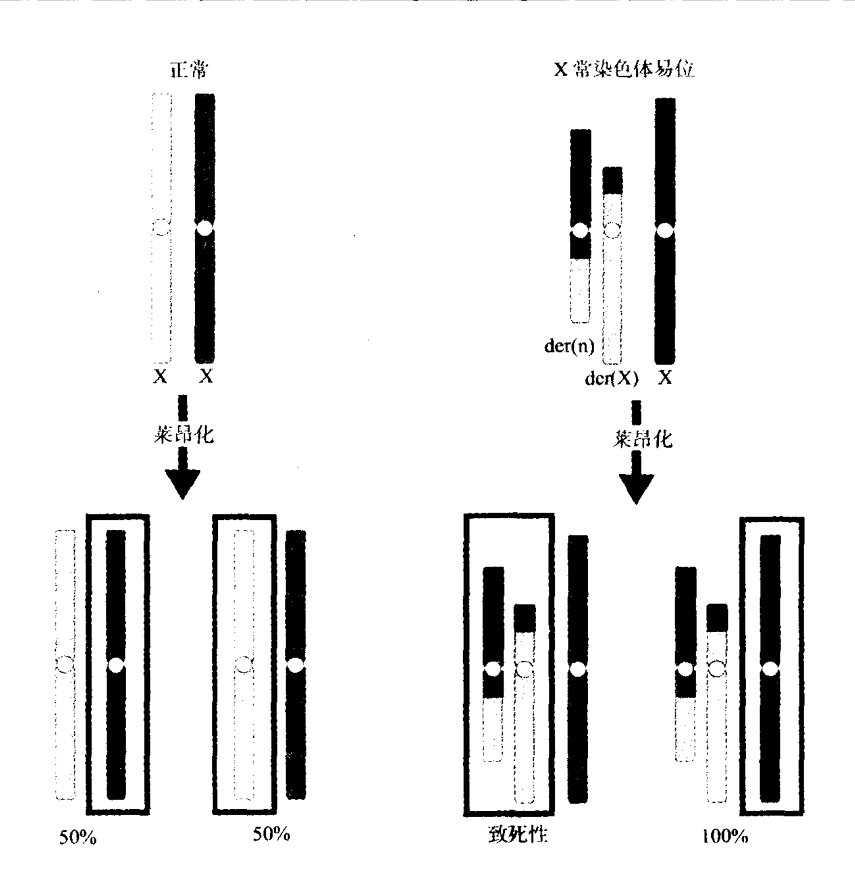


图 9-18 X 染色体与常染色体易位时 X 染色体失活的结果。正常情况下，约 50% 细胞各失活一条 X 染色体(失活用染色体周围的封闭框表示)。然而，存在 X 常染色体易位时，使易位 X 染色体失活的细胞同时也使相连的常染色体物质失活，而这对于细胞来说可能是致死性的。结果，几乎 100% 这样个体的体细胞中正常的 X 染色体失活。这解释了在 Xp21 发生 X 常染色体易位的女性即使在她的正常 X 染色体上携带正常的 *DMD* 基因，也会发生 DMD 的原因。为简便起见，图中仅表示了 X 染色体和易位的常染色体。

DNA 连在一起。该区域与 13 号、14 号、15 号和 22 号染色体短臂一样，含有核糖体基因，即编码核糖体 RNA 的重复序列。因为那些序列是已知的，所以克隆该女患者连接核糖体序列和 X 染色体的 DNA 断裂点已成为可能。正如期望的一样，这段 X 染色体序列位于 DMD 基因内。

另一研究组使用有 Xp21.2 微小缺失的男孩的 DNA(图 9-19)。利用一种机敏的消减技术，他们获得了在正常 X 染色体存在而在这个男孩 DNA 中缺乏的 DNA 克隆片段。如图 9-11 中的一个这样的克隆片段 pERT87，用于 DNA 印迹杂交；46，Y，del(X)泳道为有 Xp21.2 缺失的男孩的 DNA。当他们检测未见细胞遗传学缺失的 *DMD* 男孩的这些克隆时，其中一些男患者的 DNA 中也发现了 pERT87 缺失(图 9-20)。这提示该 DNA 片段可能包括部分 *DMD* 基因自身。这个 DNA 克隆片段扩展大约包括 140 000bp，该区域探针检出

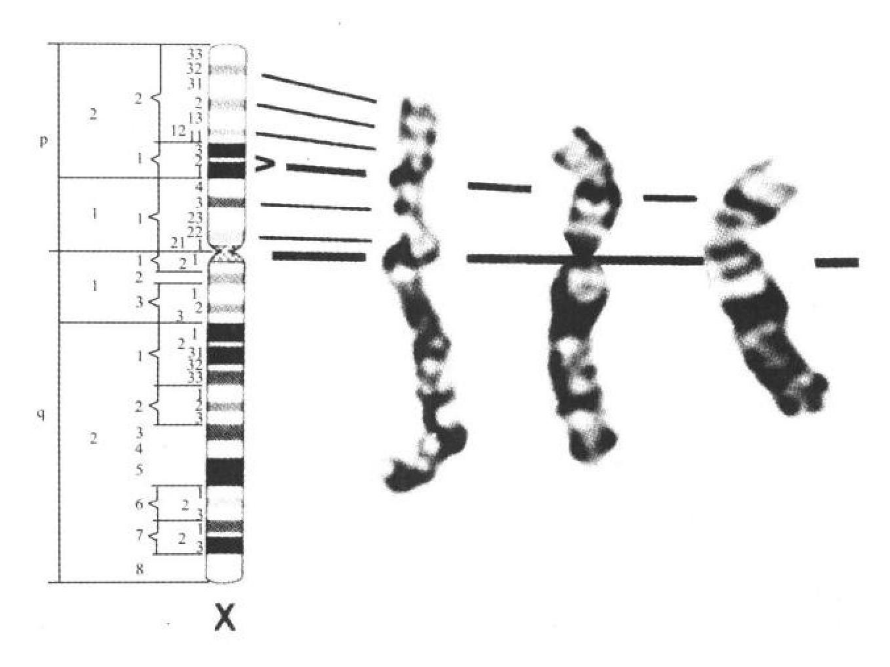


图 9-19 一名患有 DMD、慢性肉芽肿及另外两种 X 连锁疾病的男孩存在 Xp21.2 的微小臂内缺失。左侧为正常的显带模式，右侧为该男孩 X 染色体的三种代表，均存在 Xp21 区域非常微小的减少，而且均未看到明带 Xp21.2（经 Dr. U. Francke 允许引自 Howard Hughes Medical Institute at Stanford University.）。

DMD 男患在此区域两侧的缺失，说明 *DMD* 基因很大。接着，检测该区域的几个片段，观察是否含有 *DMD* 基因的外显子，最后，通过 *RNA* 印迹杂交发现了一个 16kb 的肌肉特异性转录物（图 9-21）。随后克隆该基因的 cDNA，结果发现编码 *DMD* 基因的基因组序列的确很大。作为目前已知的最大的人类基因，*DMD* 基因包含 79 个外显子，约占 2.3×10^6 bp，或占 X 染色体 1% 以上。

DMD 基因编码的蛋白质称为抗肌萎缩蛋白，是一个非常大的蛋白质（427kDa），参与肌肉细胞的收缩。抗肌萎缩蛋白的生化研究发现了若干与其相互作用的细胞骨架蛋白，称为抗肌萎缩蛋白相关蛋白复合物（图 9-22）。最近发现这些基因中几个基因突变可以造成常染色体形式的肌肉萎缩。*DMD* 基因的克隆也揭示了一种轻型 X 连锁肌肉萎缩，即 Becker 肌肉萎缩（BMD）的分子基础。Becker 肌肉萎缩也是由 *DMD* 基因突变所致。然而，BMD 通常是由于维持翻译可读框的突变或缺失引起的，导致部分抗肌萎缩蛋白的产生（图 9-23），但是数量减少或大小改变。重型 *BMD* 突变通常造成无抗肌萎缩蛋白表达或者产生截短的或者无功能的蛋白（图 9-23）。

大约 2/3 的 DMD 和 BMD 患者 *DMD* 基因有一个或多个外显子缺失。尽管整个抗肌萎缩蛋白基因的测序一般尚不可进行，但已开发了一些技术，如多重 PCR（图 9-24），用于筛查该基因内的缺失。应用免疫标记技术，分析肌肉活检标本的抗肌萎缩蛋白也成为一种有用的诊断技术，如图 9-23 所示。

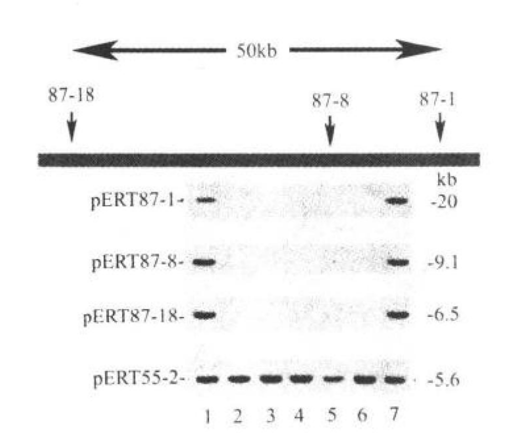


图 9-20 以 pERT87 区域的克隆为探针,7 名 DMD 男孩 DNA 的 DNA 印迹杂交。图的顶部表示了 87-18,87-8 和 87-1 探针的相对位置。1 和 7 泳道的 DNA 没有缺失证据,而 2~6 泳道均表现为 87-1、87-8 和 87-18 探针的缺失。这些男孩在细胞遗传学检测均未发现缺失(经 Monaco AP, Bertelson CJ, Middlesworth W 等允许引自 Deletion of deletions spanning the Duchenne muscular systrophy locus using a tightly linked DNA segment. Nature 1985; 316: 842~845.)。

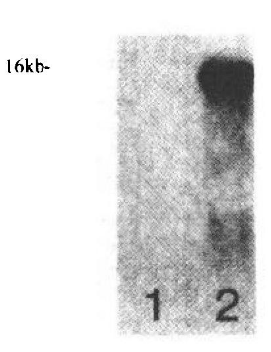


图 9-21 RNA 印迹杂交示取自胎儿骨骼肌的 16kb RNA 转录物(泳道 2),但不存在于非肌肉细胞 RNA 样品中(泳道 1)。探针是从 pERT87 区域的克隆中经 cDNA 文库筛选得到的 cDNA 克隆,这是一个很大的 mRNA,代表了正常的 DMD 基因转录物(经 Monaco AP, Neve RI, Colletli-Feener C 等的允许引自 Isolation of canadidate cDNAs for ortions of the Duehenne muscular dystrophy gene. Nature 1986;323:645~646.)。

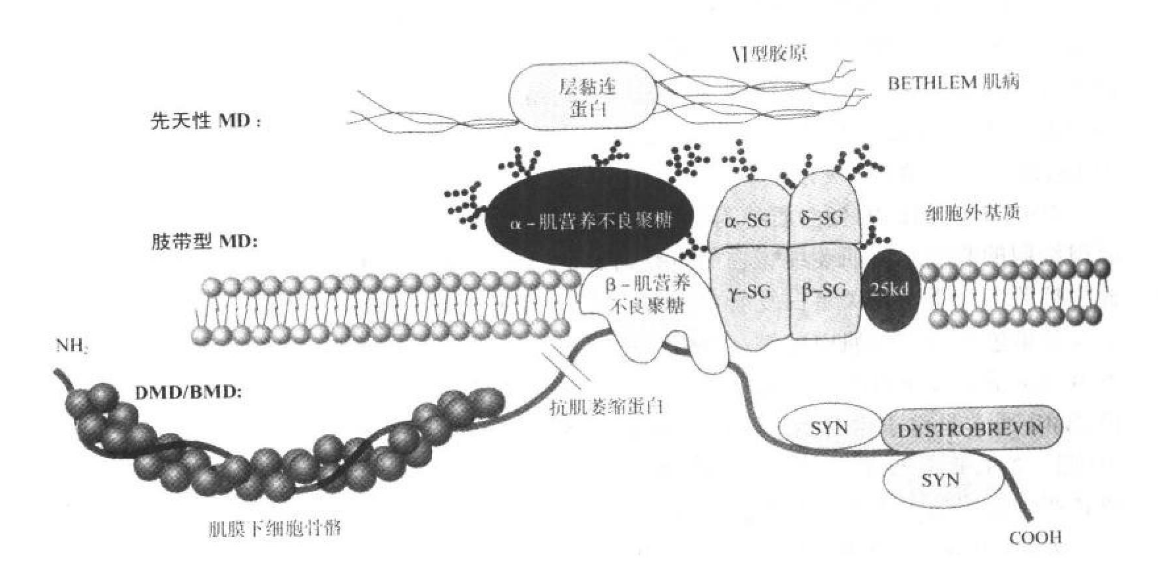


图 9-22 抗肌萎缩蛋白和抗肌萎缩蛋白相关蛋白复合物结合模式图。抗肌萎缩蛋白以粗黑线表示。其 N 末端与肌内膜下细胞骨架相作用,C 末端与 β 肌营养不良聚糖相互作用,使抗肌萎缩蛋白和细胞外基质蛋白层黏连蛋白连在一起。其他相关蛋白包括 sarcoglycans(SG)、syntrophins(SYN)和 dystrobrevin。Sarcoglycan、层黏连蛋白和Ⅵ型胶原发生突变可以分别导致常染色体肢带型肌营养不良、一种先天性肌营养不良和 Bethlem 肌病(引自 Courtery of Dr. J. S. Chanberlian University of Michigan.)。

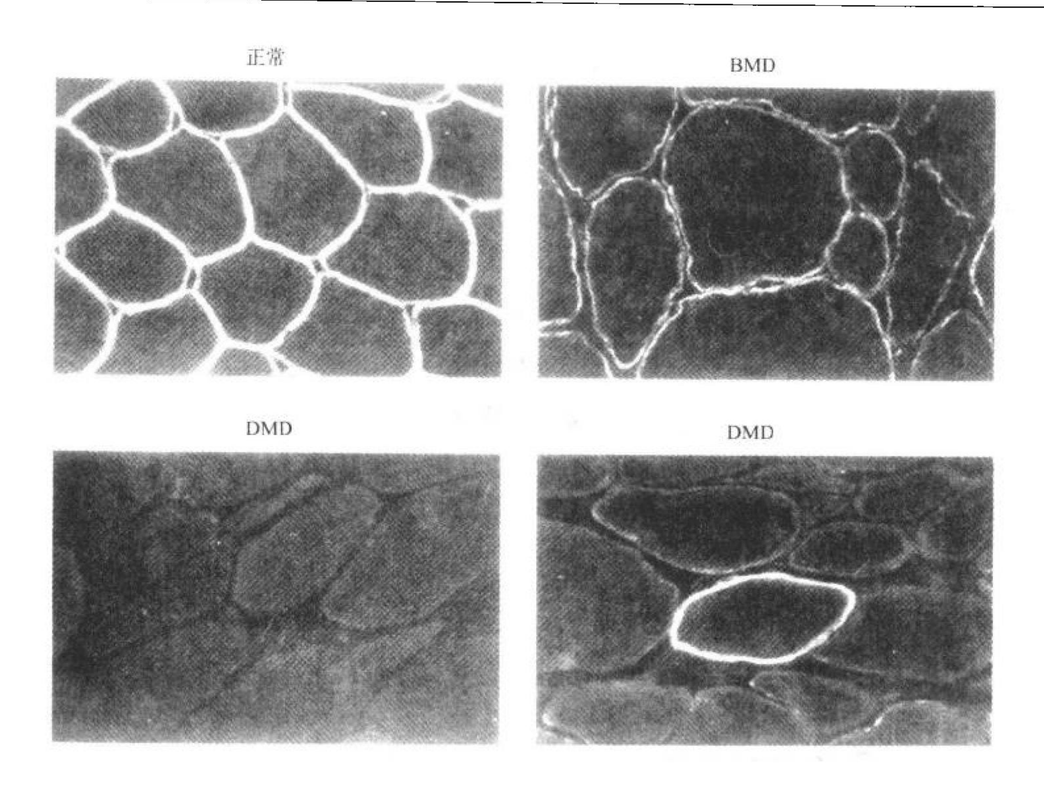


图 9-23　免疫荧光检测抗肌萎缩蛋白。应用免疫荧光技术，用抗肌萎缩蛋白单克隆抗体标记正常对照组和 BMD、DMD 患者肌肉冰冻切片。左上方所见的正常肌纤维四周浓染在左下方 DMD 患者肌纤维中完全缺失，而在右上方 BMD 患者肌肉中抗肌萎缩蛋白明显减少。注意在右下方 DMD 患者肌肉切片中可见一个抗肌萎缩蛋白阳性肌纤维。这种不经常出现的阳性纤维偶尔见于 DMD 患者肌肉中，被认为是罕见的“回复突变”，即二次体细胞突变已部分或全部恢复了蛋白编码序列（经 Nicholson LV，Danison K，Johnson MA 等允许引自 Dystrophin in skeletal muscle Ⅱ Immunoreactivity in patients with Xp21 muscular dystrophy. J Neurol Sci 1989;94:137 ~ 146.）。

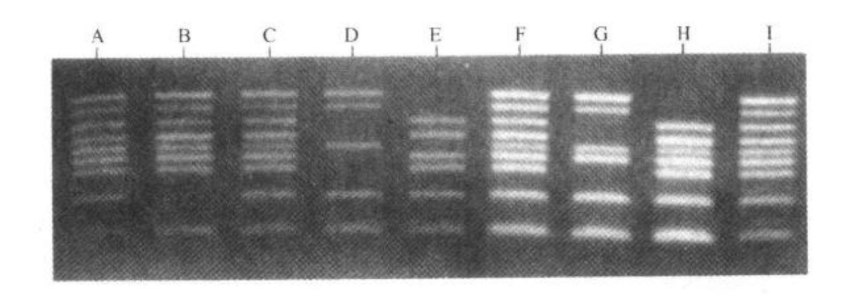


图 9-24　多重 PCR 筛查抗肌萎缩蛋白基因缺失。在这个诊断试验中，用跨越抗肌萎缩蛋白基因的 9 个不同片段的引物对一患者 DNA 样品在一个 PCR 反应中同时进行扩增。9 个 DMD 患者的产物分别在泳道 A-I 中分离。9 个扩增片段从顶部到底部分别对应外显子 45、48、19、17、51、8、12、44 和 4。患者 A、C、F 和 I 的所有 9 条带均未发现缺失。B 泳道检出跨越外显子 44 的一个缺失，D 泳道检出外显子 8、12、17 和 19 的缺失。除外显子 8 仍旧存在外，G 泳道的缺失与 D 泳道相似。E 泳道和 H 泳道均为外显子 45 和 48 的缺失，E 泳道还涉及外显子 51 的缺失。这种多重 PCR 分析能检测出所有 DMD 和 BMD 患者中约有一半的缺失（经 Chamberlian JS，Chanlerlain JR，Frenwicl RG 等允许引自 Diagnosis of Duchenne and Becler muscular dystrophies by polymerase chain reaction : a mullicenter study . JAMA, 1992;287:2609 ~ 2615.）。

(二) Huntington 病

Huntington 病是一种常染色体显性疾病,约为 20 000 人中累及 1 人。尽管外显率很高(80 岁时为 100%),但平均直到 37 岁才发病。症状主要包括个性改变、记忆丧失和一

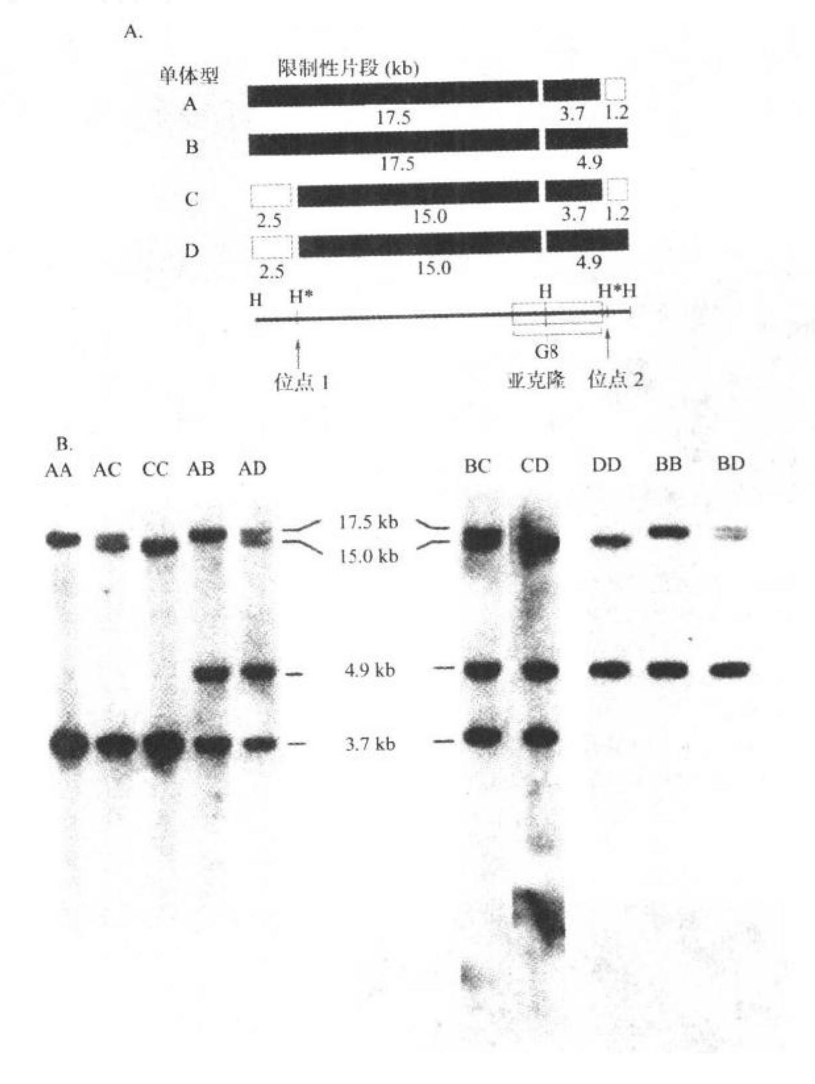


图 9-25 G8 限制性片段长度多态分析。A. G8 基因座示意图。这个 RFLP 比我们所描述的其他 RFLP 更复杂,因为它同时检测 *Hin*dⅢ两个多态位点。多态位点以星号标示,也列出了非变异位点。根据位点 1 或 2 是否存在,每条染色体可出现四种不同的形式,分别为单体型 A、B、C、和 D。探针的位置标在图的底部。值得注意的是因为探针未覆盖这些片段,所以使用这个探针进行 DNA 印迹杂交时无法观察到 C、D 单体型的 2.5kb 片段和 A、C 单体型的 1.2kb 片段。B. DNA 印迹杂交显示 *G8* 基因座 10 种可能的基因型。注意观察到的 *AD* 基因型和 *BC* 基因型的模式相同。因此,只能通过分析近亲加以辨别(经 Gusella JF, Tanzi RE, Anderson MA 等允许引自 DNA markers for nervous system diseases. Science 1984;25:1320 ~ 1326.)。

系列特殊的运动障碍,如舞蹈病(四肢的不自主运动)。病理诊断发现,该病死者均有部分脑基底神经节处明显的神经元缺失,尤其是尾状核部分。尽管过去进行了多年的研究,通过比较 Huntington 病患者和正常人的组织,仍未发现任何特殊的蛋白质异常。因此,没有任何构成 Huntington 病基础的遗传缺陷线索可进行功能克隆。同时,在任何患者均未发现特定区域的染色体异常。

Huntington 病是第一个通过系统筛查多态性遗传标记,进行连锁分析而定位的遗传病。这种方法通常称为基因组扫描,已成为标准的人类致病基因定位克隆的首要步骤。20 世纪 80 年代早期,随着对利用 DNA RFLPs 进行这类分析的认识,由于 Huntington 病是常染色体显性遗传,外显率高,并容易获得大家系的特点而成为首选研究目标。因此,研究人员开始大量收集家系资料和开发一系列探针,希望在发现某种连锁前试用 100 ~ 300 个探针。幸运的是,试用的第 12 个探针,称为 *G8* 与 Huntington 病基因座(HD)紧密连锁。这次幸运的发现立即推动了其他研究人员进入利用 DNA 多态进行连锁分析的研究领域。如图 9-25 所示,*G8* 标记可以检出*Hin*dⅢ的两个多态位点。因为存在两个多态位点,就有四种可能的染色体模式(单体型),根据位点是否存在命名为 A、B、C、和 D。图 9-25 中列出了 Southern 印迹杂交已观察到的所有可能的基因型。

图 9-26 表示了 Huntington 病一个很大的委内瑞拉家系的一部分。由于这个家系的受累个体数多,所以非常有利于进行连锁分析。家系中列出了 RFLP 分析得到的 *G8* 基因座的基因型。检查家系揭示所有 Huntington 病患者均从其患病双亲遗传了 *C* 等位基因,Ⅳ代中的一个以阴影表示的个体除外,该个体必定是一个重组体。再请注意,RFLP 分析仅仅帮助人们在该家系中追踪 Huntington 染色体,有 *C* 等位基因并不等于患有 Huntington 病,该家系中几个无亲缘关系的配偶也有 *C* 等位基因的事实说明了这一点。从这个家系及其他家系的计算所得的 LOD 值表明 *G8* 位点和 Huntington 病基因座间的遗传距离大约 3cM。综观全世界正被研究的其他家系,G8 和 HD 之间的连锁关系是毫无问题的,支持连

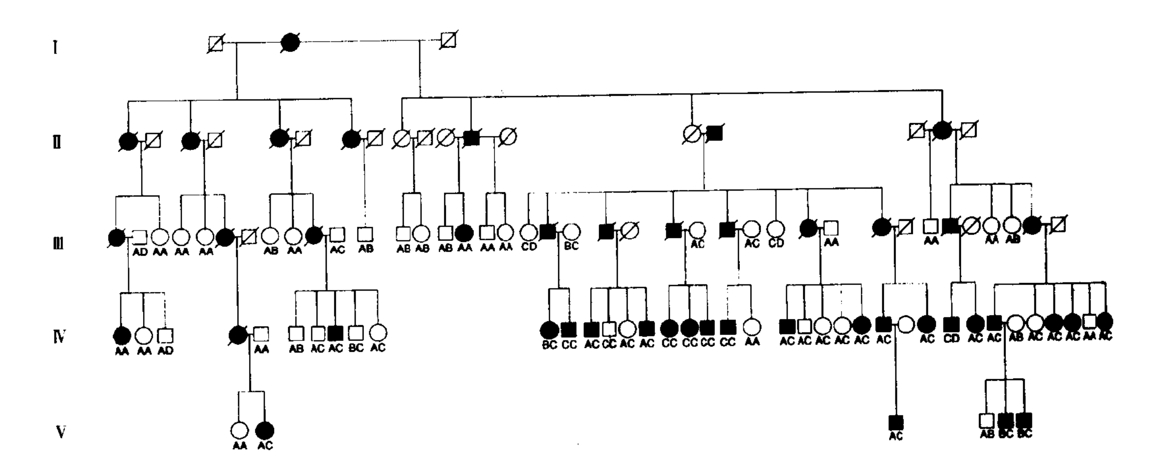


图 9-26 一个 Huntington 病委内瑞拉大家系的一部分。注意前几代中的许多人已经死亡,然而用计算机程序依据孟德尔式遗传推断死者的所有可能的基因型仍可以进行连锁分析。注意,除了第Ⅳ代中有阴影个体外,所有受累个体在 *G8* 基因座均含有 *C* 等位基因(经 Gusella JF, Tanzi RE, Anderson MA 等允许引自。DNA markers for nervous system diseases. Science 1984;225:1320 ~ 1326.)。

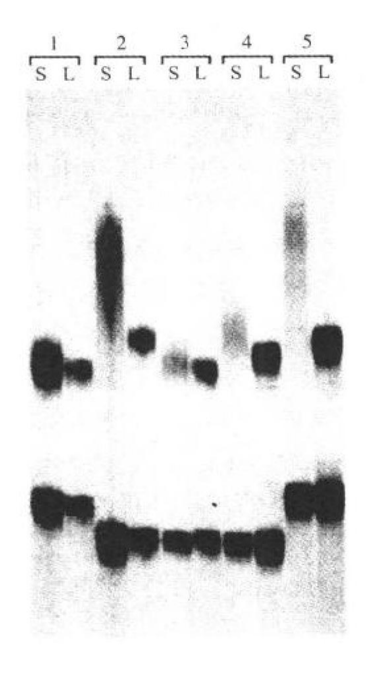


图 9-27 *HD* 基因三联体动态重复的 PCR 分析。设计 PCR 引物特异性扩增 *HD* 基因 5'端含 CAG 重复编码序列的基因片段。DNA 取自图 9-26 中委内瑞拉 HD 家系的 5 个成员。"S"泳道用取自精子的 DNA,"L"泳道为淋巴母细胞 DNA。每个泳道内较低的条带来自正常染色体,含 19 ~ 24 个 CAG 重复拷贝。每一个泳道内较高的条带为带有 *HD* 基因突变的染色体的产物。在测量淋巴母细胞 DNA 泳道时,这些扩增的等位基因长度范围含有的三联体重复拷贝数从 45(1-L 泳道)到 52(2-L 和 5-L 泳道)。因此,所有 5 个个体(研究时年龄 24 ~ 30 岁)均携带 *HD* 基因突变,预计可发展成 Huntington 病。注意从精子 DNA 样品扩增的等位基因大小范围通常大于从淋巴母细胞 DNA 扩增相同的等位基因。这是由于在男性减数分裂过程中该三联体重复非常不稳定,常导致 *HD* 等位基因扩增,可解释当疾病由父方遗传时临床发现的 HD 早期发作。实际上每一个正常等位基因(较低的条带)含有一簇距离密切的条带,这是由于 PCR 扩增过程中 DNA 聚合酶"滑动"造成的人为现象(经 Duyao M, Ambrose C, Myers R 等允许引自 Trinucleotide repeat length instability and age of onset in Huntington disease. Nat Genet 1993;387 ~ 392.)。

锁的机会大于 1080:1(LOD 值为 80)。如此高的 LOD 值只有研究许多家系中大量患者后才能获得。*G8* 标记位于 4 号染色体短臂顶端,说明 *HD* 位点也一定接近该位置。

尽管 *G8* 和 *HD* 间的连锁关系在 1983 年首先报道,完成定位克隆过程到鉴定 Huntington 病致病基因本身和 HD 病遗传缺陷的特征描述,又花费了十年。然而,一旦 G8 及其他多态标记与 *HD* 紧密连锁得到证实,就可以对 HD 高危人群,在其出现症状前和用标准医学方法可能进行诊断很早之前,就可以进行 DNA 预测。这种症状前 DNA 诊断方法带来了一系列新的复杂的临床和伦理问题,其中一些问题在第十二章和十四章中讨论。

1993 年克隆 *HD* 基因使有关基因突变的鉴定成为可能。描述 G8 附近基因的特性,检出患者区别于正常人的遗传学改变,这是多年工作的最终成果。Huntington 病是由 HD 蛋白氨基末端三核苷酸动态重复引起的,称为 Huntingtin(见第八章图 8-44)。尽管 Huntingtin 蛋白的功能尚不清楚,*HD* 基因的克隆为最终认清 HD 的病理生理学和开发新的治疗方案提供了强有力的研究新手段。

如图 9-27 所示,HD 三联体动态重复可通过 *HD* 基因座的 PCR 扩增直接检测。CAG 重复范围在正常的染色体为 10 ~ 30 拷贝,在 HD 染色体为 36 ~ 121 拷贝。检测 *HD* 基因中 CAG 重复的长度就可以直接诊断 HD,不像连锁分析那样要求必须检测其他家庭成员,也不必担心可能出现重组问题。CAG 重复的长度与症状出现的年龄一般相关(重复的长度越长,预示出现症状越早),但这一点对于临床应用不够精确。

HD 分子诊断的发展为其他成年后发病疾病的症状前检测的一般方法提供了有价值的模式。最近随着引起如乳腺癌和结肠癌这些常见病的高发基因突变的检出,由 HD 的 DNA 检测引发的复杂伦理问题更受到广泛关注(见第十二和十四章)。

(三) 囊性纤维化

囊性纤维化(CF)是白种人中最常见的常染色体隐性遗传疾病之一,新生儿中约1/2500受累。从 Hardy-Weinberg 平衡定律可计算出杂合携带者频率约为 1/25。该病的临床特征为肺和胰腺的黏稠黏液分泌。出生时可由一种小肠梗阻,称为“胎粪样肠梗阻”的情况而识别。通常通过汗液中氯化物水平升高而作出诊断,这是 CF 既高度敏感又具特异性的表现。该病患儿通常有胰腺功能不全,补充酶类可以控制。但患儿的慢性肺感染更难以控制,尽管应用抗生素和物理疗法对胸部感染进行积极治疗可提高存活年限,但目前的平均存活年限为 30 年左右。虽然收集的大量证据显示其跨膜离子运输过程中存在异常,但几十年的研究未能检出引起囊性纤维化的特异基因产物。

由于 CF 的重要临床意义和功能克隆的失败,它也成为应用连锁分析进行基因定位的主要目标。1985 年首次说明了 *CF* 与 7 号染色体的遗传标记连锁。随着 7 号染色体其

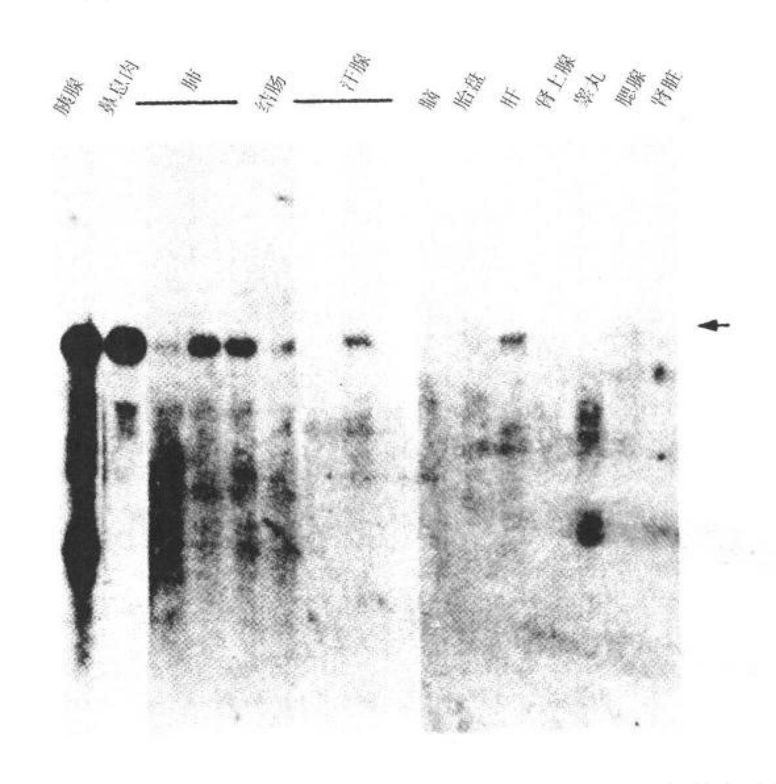


图 9-28 RNA 印迹杂交显示囊性纤维化基因表达模式。印迹杂交的探针由正常个体汗腺的 cDNA 片段制备。每一条泳道有不同的正常人组织的 RNA。一个 6.2kb 大小的 mRNA 在肺、胰脏和汗腺中明显可见,而未见于其他组织(如脑组织)。这种所见与预想的 *CF* 基因转录物一致,据推测它在上述器官中发挥正常的功能。当该基因的两个拷贝都有缺陷时,就产生囊性纤维化(经 Riordan JR, Rommens JM, Kerem B 等允许引自 Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA。Science 1989;245:1066~1073.)。

他遗传标记的检出，人们发现 *CF* 基因与遗传标记 MET 和 J3.11 距离很近。物理制图技术显示这些标记之间的距离约为 1.6Mb，是足可包括约 50 个基因的大小的区间。随后又在这段区域中得到其他的遗传标记，使 *CF* 基因定位范围进一步缩小。

与前文所述血色素沉着症和 HLA 相似，随着多态标记距 *CF* 基因越来越近，人们观察到连锁不平衡程度亦增加。在这种以及相似的情况下，连锁不平衡的程度有时可为该基因的实际定位提供有用的信息。超过 70% 的囊性纤维化患者染色体中可观察到的有

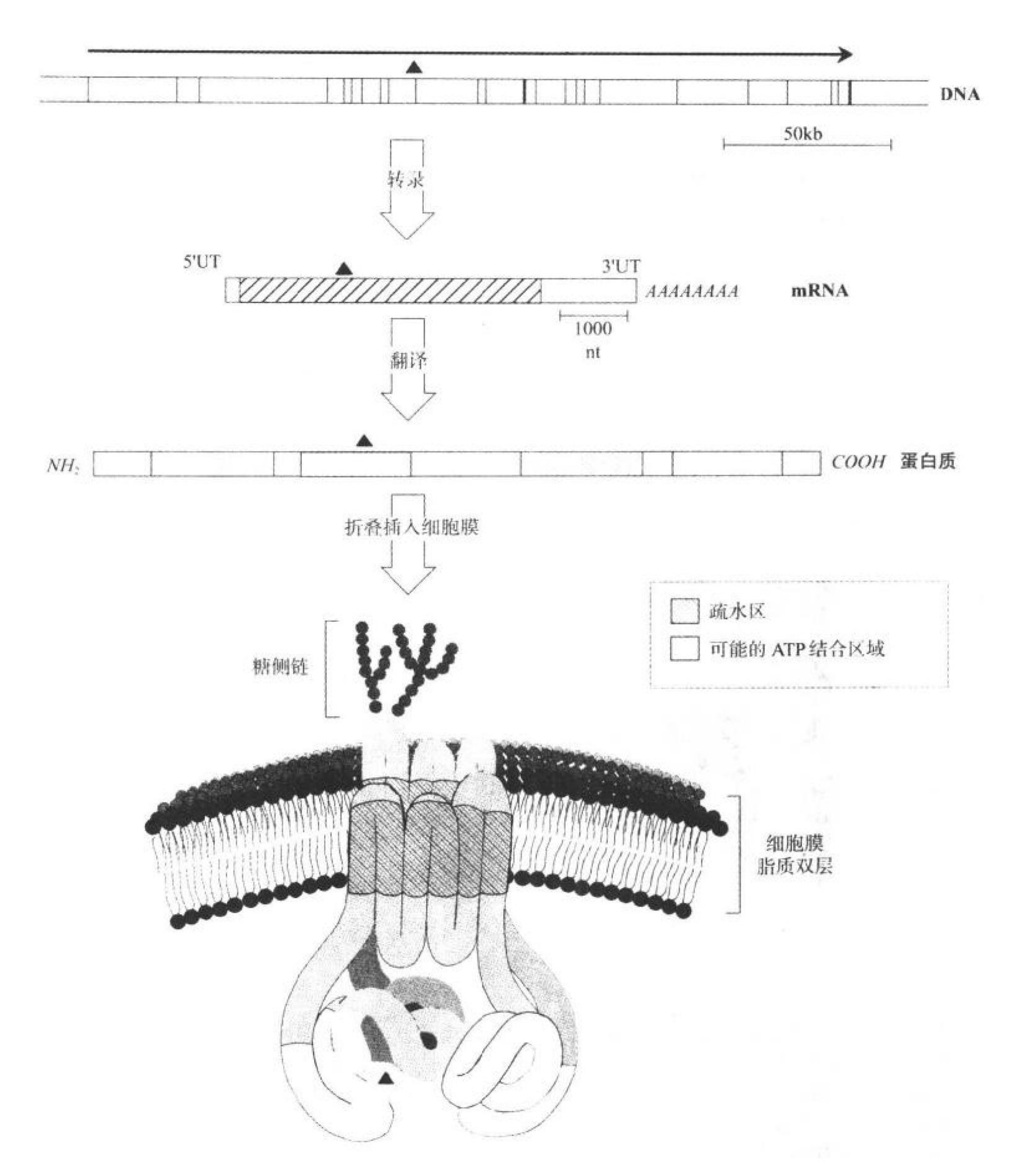


图 9-29　定位克隆鉴定囊性纤维化相关的基因、转录物和蛋白。根据与其他已知蛋白的相似性，此处所示 CFTR 折叠结构是假想的。黑三角表示第 10 外显子 3 个碱基缺失的位置，代表囊性纤维化常见的突变。这导致蛋白产物丢失一个氨基酸，用黑三角表示。

特定模式的多态等位基因,仅见于约25%正常个体的染色体中。

人们通过遗传分析认真研究了约500 000bp的候选区域,最终克隆了 *CF* 基因。在研究DMD时,重要重排是非常有用的,而在研究 *CF* 基因时,常无可用的重要重排,这正是检出正确 *CF* 基因的主要问题。事实上,在真正的 *CF* 基因最终克隆之前,曾经在候选区域中研究过三个可能的基因,但发现它们和 *CF* 基因无关。

如图9-28所示,应用第四个候选基因作为探针进行RNA杂交,在肺脏、胰腺、汗腺、肠和肝脏都检出一种转录物,所有均在CF中受累。该基因很大,长250 000bp,含26个外显子(图9-29)。mRNA转录物长6129bp,编码一种1480氨基酸蛋白质。该蛋白的序列被称为CFTR(囊性纤维化跨膜调节作用因子),与其他已知蛋白相似,它锚着于细胞膜中,与ATP结合。以后的实验中证实了这些假想,而且也确证了该蛋白具有氯化物通道的作用。

为了证明这个候选基因的正确性,有必要鉴定与囊性纤维化相关的基因突变。从囊性纤维化患者汗腺中克隆并测序的cDNA结果揭示,在第10外显子有一个异于正常序列的明确差别:发现一例三碱基缺失。这种缺失会引起第508位氨基酸(苯丙氨酸)的丢失(图9-30),这代表了一种导致囊性纤维化的突变而不是中性多态,并在随后对携带正常序列或密码子508缺失(ΔF508)序列的重组蛋白进行功能分析时得到证实。值得一提的是这种ΔF508可见于70%的囊性纤维化染色体,并且几乎出现于所有先前显示带有 *CF* 连锁不平衡的常见多态单体型囊性纤维化染色体中(见上述)。因此,ΔF508是在 *CF* 基因克隆之前第一个由连锁分析提示的祖先或建立者突变。

正常

DNA	*..GAA AAT ATC ATC TTT GGT GTT TCC..*
蛋白	Glu Asn Ile Ile Phe Gly Val Ser
位置	504 505 506 507 508 509 510 511

囊性纤维化

DNA	*..GAA AAT ATC AT- --T GGT GTT TCC..*
蛋白	Glu Asn Ile Ile Gly Val Ser

图9-30 导致 *CF* 基因的常见突变。外显子10CTT缺失引起1480个氨基酸的蛋白质第508密码子(苯丙氨酸)丢失。而第507位的异亮氨酸并没有发生改变,因为ATC和ATT都编码异亮氨酸。

CF 基因克隆的成功已获得的主要结果。最直接的是,它使便捷检测这种常染色体隐性遗传性疾病的杂合子(4%白种人)成为可能,进而使一些有生育患儿风险的夫妇在建立家庭之前得到这一信息(详见第十二章遗传筛查)。如上所述约70%的 *CF* 等位基因携带ΔF508,通过对大量患者的研究,又检出了600多种其他类型的突变。如第十二章所述,应用70种最常见 *CF* 突变嵌板对CF携带者进行DNA检测筛查,可在北欧白色人种中检出90%的携带者,在其他选定群体中检出60%~98%的携带者。新近有专家组建议对所有准备和已经怀孕的夫妇进行 *CF* 携带者的筛查工作。

第二,*CF* 基因克隆为理解该病的生物学已取得重要的成果。基础研究的不断进展能导致药物和基因治疗领域产生新的方法(见十三章)。最后,即使没有遗传基因重排引路,定位克隆鉴定 *CF* 基因的成功证实了这种方法的可行性,从而为大量正常基因功能的尚无信息的人类疾病基因的鉴定提供了方法。

六、定位克隆方法的概述

随着分子遗传学技术的不断进展,用定位克隆确定人类疾病的致病基因已日益常规化。总体策略图解见图 9-31。第一步一般用连锁分析方法对感兴趣的基因制图定位。有时一种间质缺失或染色体易位可提供第一线索,如 DMD 的例子中,然而现在许多不首先借助细胞遗传学定位致病基因的情况越发增多。此时,需要进行全基因组扫描以寻找与已知位点的隐匿 DNA 连锁的标记。实际上,识别足够并可为连锁分析提供具有统计学意义的受累家系经常是一个限速的步骤。现在连锁分析中的许多步骤都已经自动化,成千上万的样本可以在一天内全部完成。一律用 PCR 技术进行基因分型,用强大的计算机程序来分析数据,并跨越全基因组筛查成百上千的遗传标记。连锁分析取得初步提示后,在有意义的区域内新的标记中分类以缩小候选范围。患者测试遗传标记与疾病表型之间的重组现象可以进一步缩小候选区域。

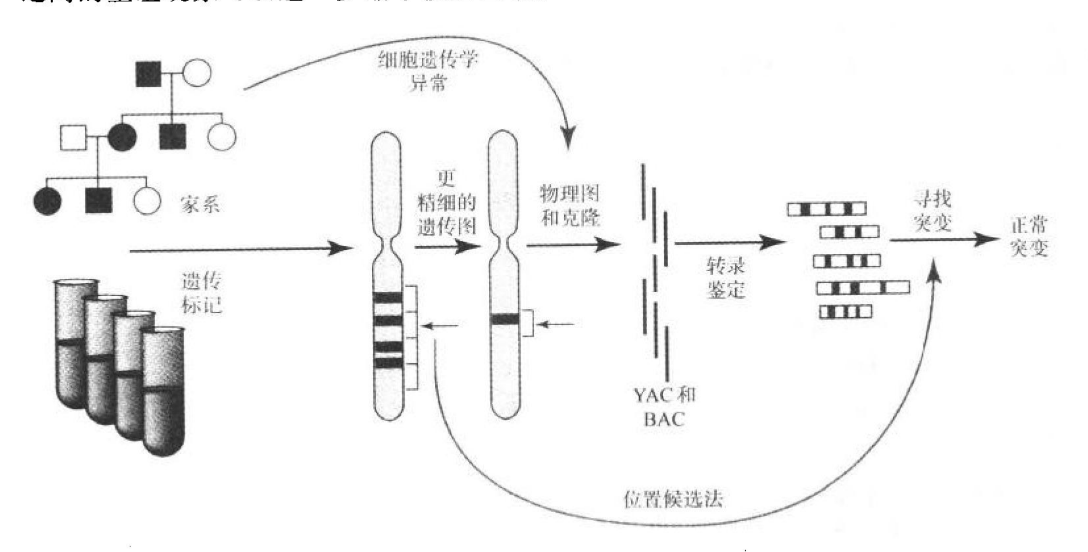


图 9-31　定位克隆常用方法简图。虽然独特的细胞遗传异常偶尔提供一条捷径,但通常是用大量的 DNA 多态标记(基因组筛查)对许多受累家系进行连锁分析初步定位致病基因。有时,根据基因的功能及其靠近最初已定位致病基因的已知位置可快速确定某一可能的候选基因(下方箭头)。这种捷径被称作位置候选方法。

分析完足量患者和遗传标记可将候选基因区域缩至 1cM 或更小之后,下一步通常是在有意义区域中确定目的基因。直至最近,这种方法主要依靠用特定载体如 YAC 和 BAC (见图 9-16),从候选区域克隆出人类的长片段 DNA。现在这一步从主要根据连锁分析信息的遗传图转为物理图,这样表达了 DNA 的真正物理位置。然后,用多种方法在这些较

长的 DNA 片段中识别潜在的基因。随着人类基因组计划的不断进展(见第十章),不断地定位人类全长基因互补序列成为可能。现在在计算机键盘上搜索日益扩展的公共基因通常可迅速地在某一基因区域检出候选基因。另外,YAC 克隆重叠的物理图现已收集 95%以上的人类基因组(见第十章)。现在从资源实验室常可获得某些特定克隆计划的相关 YAC,而不必进行过去必需的繁琐的文库筛查。在今后几年中,人类 50~100 000 个基因中的大多数将被定位于遗传图谱上,随着人类基因组计划的完成,将得到人类所有的基因序列。

在某些情况下,检验某一新确定的人类疾病基因位置附近的许多基因会快速分析检出某一很有吸引力的候选基因,缩短整个过程,避免更详细定位与大规模克隆。例如,家族性心肌肥大是一种心肌发育异常疾病,定位于 14q,靠近 α 和 β 心肌重链基因。已知这些基因是心肌的重要成分,所以它们立即成为强有力的候选基因。因一些特定突变的发现,人们很快证实这些基因是该病的致病基因。这种“位置候选基因”法正逐步取代以前定位 *CF* 基因时应用的较繁琐的定位克隆技术(如图 9-32)。正如家族性心肌肥大的例子,现在可以从人类图谱中某一疾病基因座的定位迅速转向识别这种致病基因和相应的突变本身。

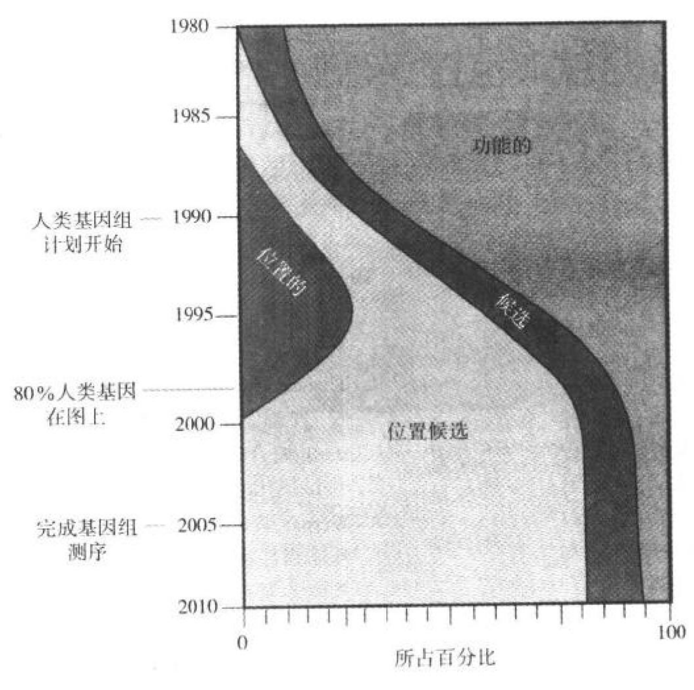


图 9-32 人类疾病基因克隆方法的演变模式。定位克隆在 20 世纪 80 年代晚期刚刚成为可能,并在人类基因组计划制图技术的推动下迅速发展起来。随着基因制图不断向更精细发展和人类全基因组测序最终完成,位置候选法的重要性将继续增加,大部分取代传统功能克隆方法。少数致病基因仍将通过单纯候选方法识别,用这种方法即使没有基因定位信息,已知或未知的基因功能常提示与某一疾病发病机理相连锁(经 Collins FS 允许引自 Positional cloning moves from perditional to traditional. Nat Genet 1995;9:347~350.)。

表 9-2 定位克隆鉴定的疾病基因

1986
慢性肉芽肿疾病
Duchenne 肌营养不良
视网膜母细胞瘤

1989
囊性纤维化

1990
Wilms 瘤
1 型神经纤维瘤
睾丸决定因子
无脉络膜病

1991
脆性 X 染色体综合征
家族性结肠息肉
Kallmann 综合征
无虹膜病

1992
强直性肌营养不良
Lowe 综合征
Norrie 综合征

1993
Menkes 病
X 连锁丙种球蛋白缺乏症
甘油激酶缺乏症
肾上腺脑白质营养不良
多发性神经纤维瘤 2 型
Huntington 病
Von Hippel-Lindau 病
脊髓小脑共济失调 1
无脑回畸形
Wilson 病
结节样硬化

1994
Mcleod 综合征
多囊性肾病 1 型
齿状核红核苍白球丘脑下部核萎缩
脆性 X 染色体 RES
软骨发育不全
Wiskott Aldrich 综合征
早发性乳腺/卵巢肿瘤(*BRCA1*)
Diastrophic 发育异常
Aarskog-Scott 综合征
先天性肾上腺发育不全
Emery-Dreifuss 肌营养不良
Machado-Joseph 病

1995
脊柱肌肉萎缩
点状软骨发育异常
肢带肌营养不良
眼睛白化病
毛细血管扩张性共济失调
Alzheimer 病(14 号染色体)
Alzheimer 病(1 号染色体)
低血磷佝偻病
遗传多发性外生骨疣
Bloom 综合征
早发性乳腺癌(*BRCA2*)

1996
Friedreich 共济失调
进行性肌痉挛性癫痫
Treacher Collins 综合征
长 QT 综合征(11 号染色体)
Barth 综合征
Simpson-Golabi-Behmel 综合征
Werner 综合征
X 连锁性视网膜炎色素病(*RP3*)
多囊性肾病 2 型
基底细胞痣综合征
X 连锁性肌管肌病
无汗性外胚层发育异常
血色病
Chédiak-Higashi 综合征
Rieger 综合征
青年成熟期发作糖尿病(12 号染色体)

到 1996 年 12 月 9 日。此表除外首先由候选基因方法或在肿瘤细胞中体细胞突变基础上鉴定的基因。

表 9-2 说明了定位克隆识别基因的步伐不断加快。人类基因新信息的爆炸导致强大

的计算机资源的开发,以便获得这些数据,以及各类数据库之间丰富的相互联系。人类孟德尔遗传网络版(OMIM,http://www.ncbi.nlm.nih.gov/Omim/)是由 Victor McKusick 博士首创,它为人类遗传疾病编目,列举所有已知人类遗传疾病。它现在包含大约 9 000 个不同遗传基因座。OMIM 还与基因组数据库(GDB, http://gbdwww.gbd.org)连接,后者拥有关于遗传标记和遗传图上已知基因位置的丰富信息资源。

七、人类复杂疾病的基因定位

本章及全书大部分都集中于经典的孟德尔式遗传病的分子遗传学。然而,如第四章中所讨论的,许多人类疾病的遗传学构成,不能归因于单基因受累,而是多基因和环境因素间复杂相互作用的结果。随着单基因病遗传连锁分析方法的巨大成功,以及人类基因组计划产生的强有力的工具,复杂遗传疾病的研究已成为当代遗传学研究的主焦点。应用这一方法我们正在研究若干常见的人类疾病,如 Alzheimer 病、常见肿瘤、动脉粥样硬化、高血压、哮喘、糖尿病。

在基因的位置或功能以前不了解时,如何着手鉴定某一复杂遗传性状的基因?这是比我们至今正讨论用标准连锁分析研究经典孟德尔遗传方式疾病困难得多的问题。由于任一个基因的作用可能轻微,通常需要研究更多的家系,更广泛地收集遗传标记。

如第四章所综述,分析大群体中某一特定等位基因与某种疾病的相关性可识别候选基因。换言之,如第四章讨论的胰岛素依赖型糖尿病,可应用同胞对分析进行全基因组扫描。这些分析中的数学和计算问题,以及这些结果的解释都是难以解决的。虽然这些类型的研究已经确定一些常见疾病的潜在易感基因座,如精神分裂症和多发性硬化,但最初的观察结果很难重复。

即使成功地定位了一个潜在的易感基因座,因为本章所述大部分定位克隆技术都不直接适用,下一步识别特异基因本身也是一个艰巨的任务。染色体重排和大片缺失不大可能遇见,相关序列改变只引起基因功能的轻微改变,也许难以证实。对于单基因病定位克隆,识别遗传标记基因座和致病基因之间重组事件来缩短基因搜索范围至可操作区域常常是一个极为关键的步骤。然而,对于复杂性疾病,由于任何一个易感基因只可望在小部分受累个体中出现,因此通常不可能识别如此清晰的重组。所以,即使已确定某一复杂疾病的候选基因座,界定其可能染色体位置的区域仍是很大的。

作为一种替代方法,许多研究者开始是集中研究遗传性纯合群体,以期在任一这样群体中主要易感基因座数目会少一些且较易于识别。另外,来自于有限数目建立者的特定等位基因也可以通过连锁不平衡来确定其存在。

最后,动物模型为复杂人类疾病的分析提供了强有力的手段。人类基因组计划(见第十章)已为一些模型有机生物提供了遗传和物理定位方法。对于小鼠,所用方法几乎和人类所用的同样广泛。由于可进行选择性交叉育种实验,这使小鼠的修饰或易感基因的精确定位常常更加容易。此外,应用小鼠可进行基因剔除或转基因实验(见第五章),从而使对候选基因进行遗传操作成为可能。经遗传修饰的小鼠也可以用于交叉育种,以检测易感基因间潜在的相互作用。这种方法可通过小鼠动脉粥样硬化模型来说明,该方

法可检测一些基因在脂类代谢、炎症和凝血中的相互作用。

应该时刻注意,在其他哺乳类动物中检出的易感基因并不一定直接与人类相关。当初步的人类连锁或相关性分析提示,人类基因组某一区域中的易感位点与正在研究中的与小鼠染色体上的基因在进化上相关时,小鼠模型尤为重要。这种基因共同存在于同一染色体上的现象称为**同线性**(synteny),通常在相关种系的基因组的同源区中呈高度的保守性。在亚染色体区域中两物种间呈现同线性样保守性时,称为"同线性的"。例如,人类 17 号染色体的大部分区域与小鼠 11 号染色体相应区域呈同线性。现在认为同线性的保守性是应用动物模型系统进行人类遗传疾病研究的有效工具。

复杂人类疾病的遗传研究可能成为未来多年遗传研究的重要领域。易感基因的确定会使许多人类常见疾病发病风险的预测更加精确。这些分析使得大力精简高发风险个体成为可能。此外,对这些基因功能的基础生物学研究最终会导致开发新的治疗手段。

推 荐 读 物

Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic basis of inherited disease.7 th ed. New York: McGraw-Hill, 1995. Also available in an updated CD version. This outstanding text is also cited in several other chapters. The introductory section provides additional information concerning mapping and positional cloning methods. There are also authoritative reviews providing more detail for each of the specific disease examples covered in this chapter.

Ott J. Analysis of human genetic linkage. Baltimore : Johns Hopkins University Press, 1991. Textbook providing detailed mathematical consideration of genetic linkage analysis.

Collins FS. Positional cloning move from perditional to traditional cloning. Nat Genet 1995;9:347 ~ 350. Review of recent progress in positional cloning.

Lander ES, Schork NJ, Genetic dissection of complex traits. Science 1994;265:2037 ~ 2048. Overview of the challenge of genetic approaches to the analysis of complex human disease.

Collins FS. Sequencing the human genome. Hosp Pract Off Ed 1997;32:35 ~ 43, 46 ~ 49, 53 ~ 54. General review of genetic mapping and positional cloning and the role of the Human Genome Project.

附录 A:LOD 值的计算

如前所述,LOD 值是根据若两个被测基因座连锁并相距 θ 获得一组数据的似然性与若这两个基因座不连锁获得相同数据的似然比值计算得出的,LOD 值是这种优势比率的对数(lg)。

$$\mathrm{LOD} = \lg \frac{\text{两基因座连锁且相距 } \theta \text{ 所得数据的似然性}}{\text{两基因座不连锁所得数据的似然性}}$$

应用图 9-5 的系谱资料,我们可以计算出 *Eco*RⅠ RFLP 与 *NF1* 间连锁的 LOD 值。本书前文提到,该例中当 $\theta = 0$ 时,LOD 值为 $-\infty$。由于Ⅲ7 是一个重组体,如果致病基因和标记是不可分割($\theta = 0$)的话,将不会有重组现象发生。

我们现在计算一下 $\theta = 0.05$,或者说是重组率为 5% 时的 LOD 值。Ⅰ1 有一个正常和

一个突变的 *NF1* 拷贝,并有 4.7kb 和 3.0kb 的 *Eco*RⅠ RFLP 等位基因。解决问题的关键在于确定哪一个 *NF1* 拷贝与 RFLP 基因连锁。由于存在两种可能的情况,任何一种是正确的机会是 50%或 0.5。RFLP 与 *NF1* 基因连锁的总似然性等于第一种情况(突变 *NF1* 基因与 4.7kb 等位基因在同一条染色体上)似然性的 1/2 加上第二种情况(*NF1* 基因突变与 3kb 等位基因在同一条染色体上)似然性的 1/2。

让我们先来计算第一种情况的概率。在这种情况下,*NF1* 突变基因与4.7kb等位基因有 95%的机会一起遗传,而与 3kb 等位基因间只有 5%的可能性(标记基因与 *NF1* 基因之间有重组现象发生)一起遗传。同样,父方的 *NF1* 基因的正常拷贝与 3.0kb 等位基因间有 95%,而与 4.7kb 等位基因之间只有 5%的可能性一起遗传。由于后代继承父亲的正常和突变 *NF1* 基因的机会均等(50%或 0.5),因此从父亲那里得到突变基因和4.7kb等位基因的可能性为 $0.5 \times 0.95 = 0.475$,而得到突变基因和 3.0kb 等位基因的可能性为 $0.5 \times 0.05 = 0.025$。同理,从父亲那里得到正常拷贝和 3.0kb 等位基因的可能性为0.475,而得到正常拷贝和 4.7kb 等位基因的可能性为 0.025。Ⅱ代中四个个体都继承个体Ⅰ-1 突变基因和 4.7kb 等位基因可能性为 $0.475 \times 0.475 \times 0.475 \times 0.475$,四个个体继承父亲正常拷贝和 3.0kb 等位基因的可能性为($0.475 \times 0.475 \times 0.475 \times 0.475$)。Ⅲ代中三个个体继承个体Ⅱ8 的 *NF1* 突变基因和 4.7kb 等位基因的可能性为$(0.475)^3$,五个个体继承正常拷贝和 3.0kb 等位基因的可能性为$(0.475)^5$。一个个体显示为重组体(Ⅲ7),他继承了Ⅱ8 的 3.0kb 等位基因和 *NF1* 突变基因(0.025)。这样,第一种情况(*NF1*/4.7kb 等位基因)的似然性在此家系中 $\theta = 0.05$ 时的计算值是 $0.5 \times 0.475^{16} \times 0.025$。

现在,让我们来计算第二种情况(*NF1*/3.0kb 等位基因)的概率。同样,这种情况的可能性为 1/2。然而所有计算正好与上面相反。所有个体继承 *NF1* 突变基因和 4.7kb 等位基因或 *NF1* 正常基因和 3.0kb 等位基因现在代表重组(0.025),而在第一种情况下的重组体(Ⅲ7)现在是惟一的一个非重组染色体(0.475)。这样,第二种情况的似然性为 $0.5 \times 0.025^{16} \times 0.475$(与第一种情况相比非常小)。

记住,假定 *NF1* 基因和 RFLP 标记完全不连锁,似然性比值是以Ⅱ代和Ⅲ代中具有发病风险的 17 个个体该特定遗传方式观察值似然性作分母。该例中,*NF1* 和 RFLP 各自独立地遗传,所有四种组合(正常或突变 *NF1* 基因与 4.7kb 或 3.0kb 等位基因之间随机组合)发生的概率则相同(0.25)。这样,17 个个体中这种特定的遗传类型发生的似然性为 0.25^{17}。

综合所有的计算结果将得到:

$$\mathrm{LOD} = \lg \frac{\theta = 0.05\ [1/2(\text{第一种情况}) + 1/2(\text{第二种情况})]\text{时基因座连锁的似然性}}{\text{假定基因座无连锁存在时的似然性}}$$

$$= \lg \frac{1/2(0.475^{16} \times 0.025) + 1/2(0.025^{16} \times 0.475)}{0.25^{17}}$$

$$= \lg(1442.2) = 3.16$$

该值与应用 MLINK 计算机软件计算 $\theta = 0.5$ 时的 LOD 值相似见图 9-6。

现在让我们计算 $\theta = 0.1$ 时的 LOD 值。与前面的计算非常相似,只是现在继承任何一种非重组基因的概率是 $0.45(0.5 \times 0.9)$,而继承重组等位基因的概率为 $0.05(0.5 \times$

0.1)。将这些值代入似然性等式,于是计算得:

$$\mathrm{LOD}=\lg\frac{1/2(0.45^{16}\times0.05)+1/2(0.05^{16}\times0.45)}{0.25^{17}}=\lg(607)=2.78$$

注意该 LOD 值又与图 9-6 中计算机的计算完全吻合。

比较而言,如果不存在重组体,则考虑 $\theta=0$ 时此家系的 LOD 值是非常有意义的。在该例中,每个个体继承正常或突变染色体和相应的标记等位基因的概率都为 0.5,等式将变为:

$$\mathrm{LOD}=\lg\frac{1/2(0.5^{17})+1/2(0.5^{17})}{0.25^{17}}=\lg(65\ 536)=4.82$$

这个例子说明与相距 6cM 的遗传标记相比较,检测距目的基因很近甚至是目的基因本身的标记是否存在连锁,是非常有效的。与此相似,随着其他标记与目的基因不断地接近,在定位克隆的基因组筛查初期,LOD 值越来越受到广泛的重视。

第十章　人类基因组计划

不做小的计划；它们无激动人心的魅力，而且也许不会实现。要做大的计划；目标宏伟、工作上游，须知一个壮丽而合理的蓝图一旦载入史册，将永不凋谢，即使我们离开人世间很久，它仍会显示其自身永生的价值。

——Daniel H. Burnham

上述引语写自一位芝加哥的建筑师，他对近 10 年生物医学研究领域的好像是个外行。生物学和医学在 20 世纪的多数时间内由于每个研究者的创造已经繁荣起来，他们每个人都在追求自己的假设和观念，创造了研究观察的组织，将医学从本世纪早期的一种大体描述艺术转变为复杂的、以高技术和高信息为基础的现状。物理学和工程学已有它们的“大科学”计划，如高能粒子碰撞和空间探索，而生物学仍停留在“小科学”个体研究者的范围。这种最适当转变的范例就是对遗传学和医学的未来具有重要影响的人类基因组计划（Human Genome Project，HGP）的出现，即到 2005 年，通过有组织的协作努力完成全部人类 DNA 的作图和测序。这类计划具有特殊的科学目的和划时代里程碑的意义，因此从本质上不同于传统的生物医学研究模式。但实际上这个计划已由全球的许多实验室和单位以协作和创新的精神正在进行。其意义将是深远的，因此在本书增添一章以说明这一历史性努力的目的和重要性是适当的。

一、人类基因组计划的历史

在 20 世纪 70 年代早期，获得关于 DNA 和 RNA 的序列信息是一个艰难的过程。推断 15 ~ 20 个核苷酸可能需要全年的工作。由 Sanger，Maxam 和 Gilbert 开发的 DNA 测序的有效方法（为此 Sanger 和 Gilbert 获得 1980 年诺贝尔化学奖）开辟了对长达几千个 bp 的 DNA 链高效而精确的测序之门（图 5-17）。20 世纪 80 年代中期，少数科学家开始讨论大规模测序的可能性：以“家庭手工业”方式进行基因组的少量测序，由于所有参与者的效率低，缺乏质量控制，在许多领域进展缓慢，甚至在某些领域毫无进展，那么为什么不组织一个“应急项目”替代基因组小量测序，以便一次确定全部 30 亿 bp 的序列？

在遗传学领域工作的人们实际上都同意：如果全部人类基因组序列被他们掌握，那将有利于以这种方法进行研究，并能极大地加快发现疾病基因的步伐，对人类健康也有巨大的潜在利益。然而又出现了严重的异议。在 20 世纪 80 年代中期到后期，许多研究者简直不相信有足够的技术来完成这个目标，或以合理的经费来实现它。担心提出这样一个昂贵的计划可能会转移其他研究的资源而危害整个生物医学的事业。还有人认为

这项工作是如此令人厌烦又无挑战性，以至于有创新性的科学家不愿涉足。另外也有人提出哲学方面的异议，担心以这种大规模方式获得的信息将会给我们人种带来灾难性后果。

20世纪80年代后期，世界上许多国家都曾激烈地、公开地进行过这一争论。在美国，由国家科学院(National Academy of Science)召开了一流的专家会议，而技术评估署(Office of Technology Assessment)则召开另一专家会议。两个会议得出结论：如果能够以一系列重大成就来精心地组织这个计划，它是会成功的。特别重要的是，开始工作的重点应是发展更好的遗传图，而测序只应在技术发展到能够节省成本时方可完成。

1990年10月，美国官方宣布实施人类基因组计划。其他一些国家，特别是英国、法国、日本、加拿大和德国等也都在人类基因组计划作了重大的投资。因此，这一事业从一开始就是国际合作。这是十分恰当的，因为基因组本身必定是国际性的。由国际科学家小组建立的人类基因组织(Human Genome Organization, HUGO)对此计划在各国的实施进行协调工作。

二、人类基因组计划的目标

从一开始，该计划的目标和划时代的意义就很明确。图10-1以图解的形式表明人类基因组的三个主要科学目标。成功地达到所有这些目标需要一种支持，即开发更好的技

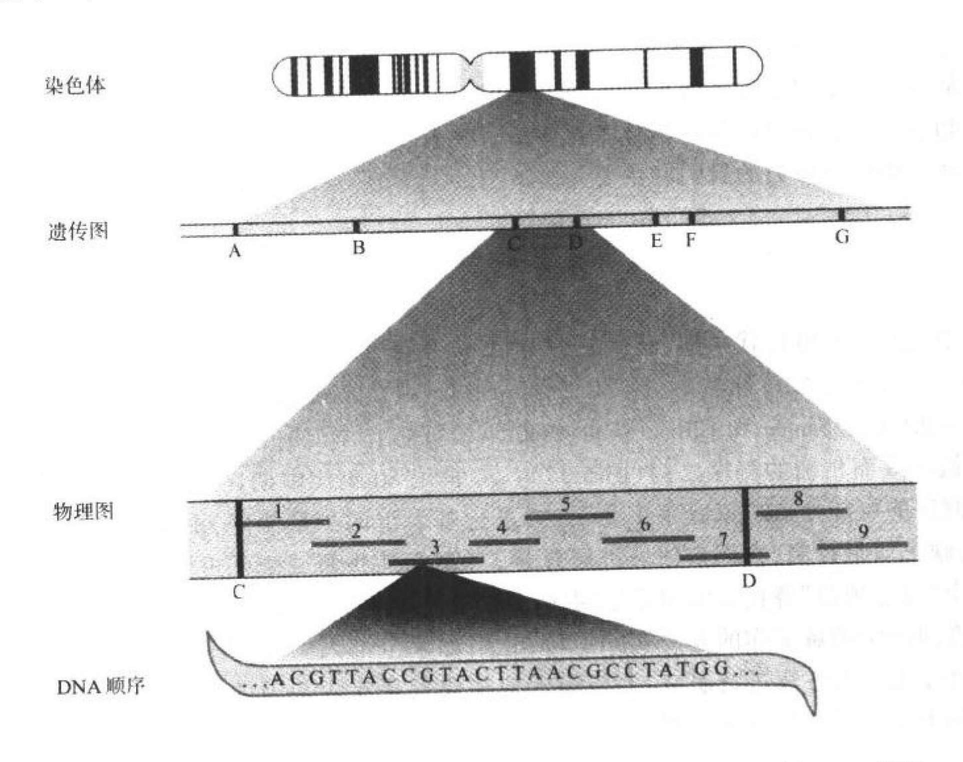


图10-1 人类基因组计划的三个主要科学目标是发展遗传图、物理图和DNA测序。在此过程中，人类基因组全部的30亿bp(大约含80 000个基因)将于2005年被确定。

术,以便能够处理因人类基因组的规模而积累的非常大量的信息。为此,正在努力推进作图和测序所必需的所有步骤的自动化。这方面需要物理学家、化学家、工程师和计算机科学家的参与,使人类基因组计划成为多学科的研究计划。这在根本上已完全解决了该计划可能的沉闷状态。一般来说,重复的任务都是自动化的,以便让科学家们自由地探索并创造性地解决 30 亿 bp(3-gigabp)基因组的问题。

(一) 遗传图

第九章中的资料概述了连锁分析和定位克隆方法鉴定人类疾病易感基因的能力,即使早先对这些基因的功能并无认识。起初开发这些方法是为了孟德尔式单基因病,而最近已扩展到多基因病。只要有一个可用而健全的人类基因组的遗传图,这些方法是会成功的,因为这是任何定位克隆计划的第一步。因此,获得由高信息量的、平均相距 2~5 厘摩(cM)的标记构成的遗传图是人类基因组计划最优先的重点。标记产物的技术学迅速地从限制性片段长度多态性(restriction fragment length polymorphisms, RFLPs)转变到简单重复序列(simple sequence repeat, SSR)微卫星(见第五章),原因在于后者的高信息含量并

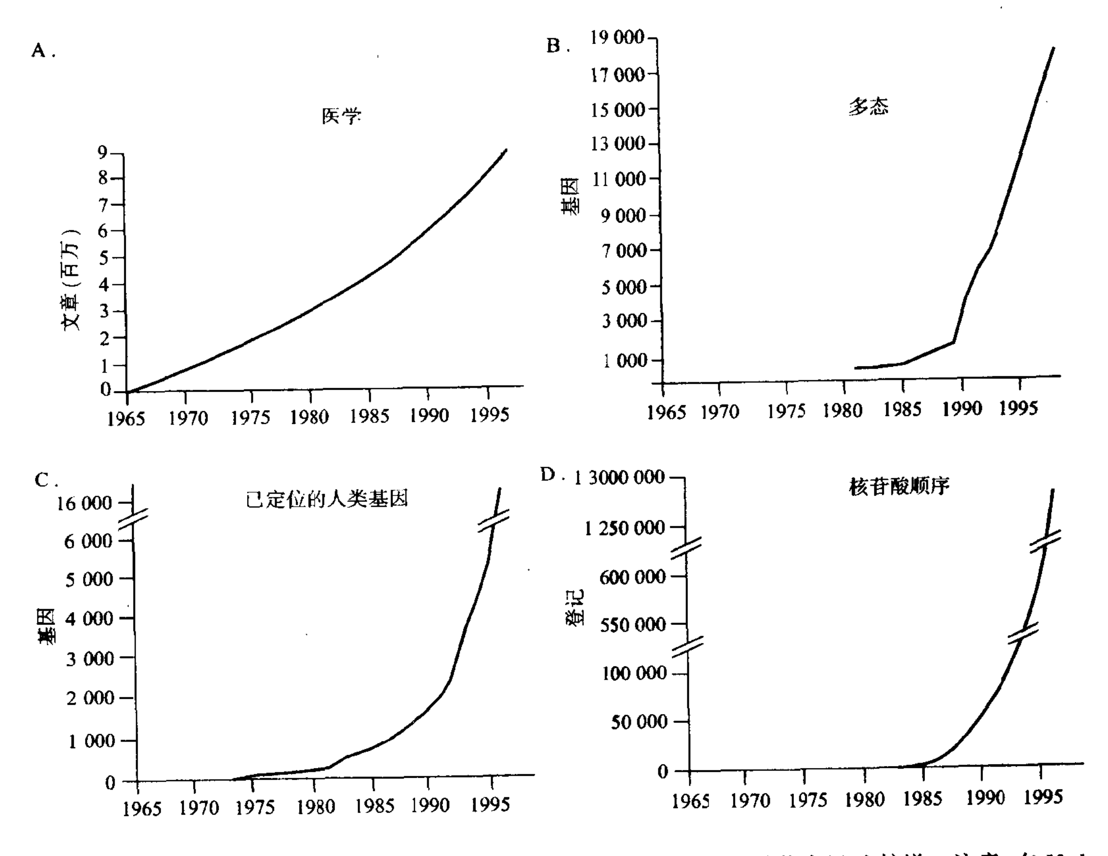


图 10-2 最近 30 年遗传学信息发生了引人注目的增长。总体上,医学信息迅速扩增。注意:在 Medline(起自 1965 年)索引的论文数目以多于线性率增长,而遗传学信息的增长更为迅速,如多态性、已定位的人类基因和 DNA 序列信息等,主要按指数增长。

易于适应聚合酶链反应。截止至 1994 年，已有 5 000 多个高信息标记定位于图上，比原先计划的平均密度要大 4～8 倍，而且提前一年多完成。图 10-2 表示遗传图的增长率和进展的其他一些参数(见下述)。图 10-3 表示 4 号染色体上标记分布的情况，表明现在可以得到高密度标记。因为有这样一大批遗传标记，倘若再有足够数量的家系用于分析，那么几乎人类任何遗传成分的表型应该都可以作图。

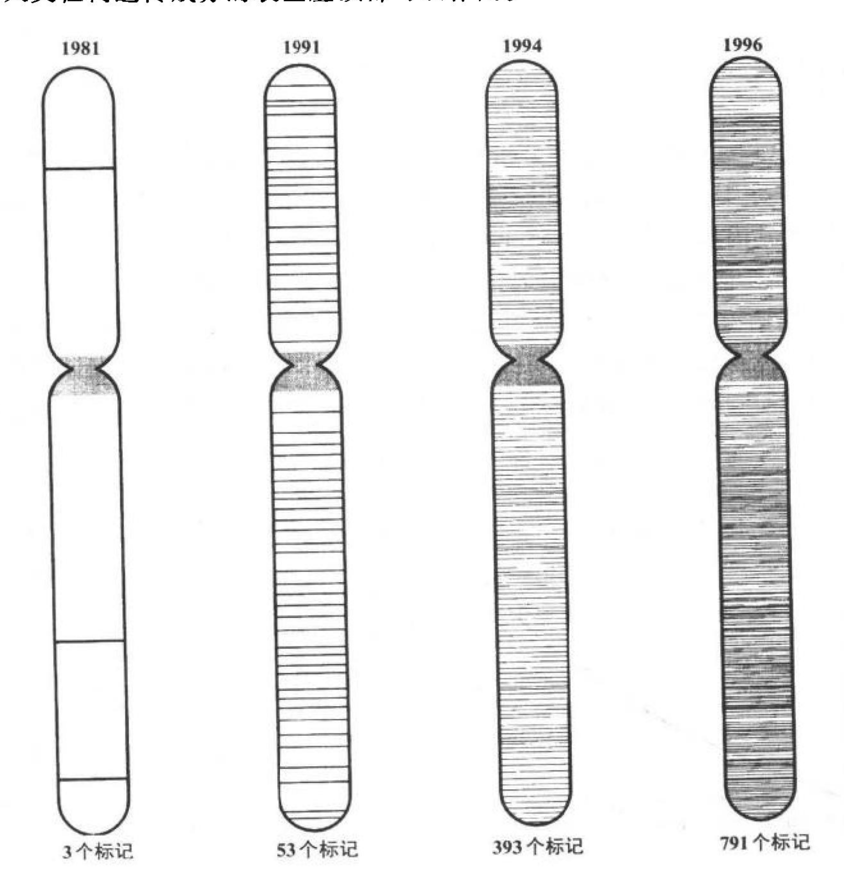


图 10-3　遗传图的迅速发展正是人类基因组计划的一个重要成果。作为一个例子，此图显示了 4 号染色体最近 15 年来可以利用的遗传标记数。在 1981 年只有 3 个限制性片段长度多态(RFLP)，而 1996 年该染色体共有 791 个标记，其中多数是微卫星，能由 PCR 获得。

(二) 物理图

连锁分析成功地将一个特异疾病基因进行一般的定位，而定位克隆者则希望追寻该基因本身。因此，正如在第九章所描述的，下一步就是“掌握”候选区域内全部 DNA，并通过它来寻找与表型可能有关的潜在基因。在 1～10 百万 bp 大小范围内产生一系列的重叠克隆片段(这在定位克隆计划中必然经常涉及)，如果不得不从起步线开始的话，这真是个令人沮丧的任务。再者，还存在经济规模的问题，即在未来的一个时期，生产这样一

个整个人类基因组的物理图比每次在一个区域内生产短片段要更为有效。毕竟,在适当的时候,人们将会发现整个基因组所含载的有意义的序列。

为了产生这样的物理图,运用可以克隆大片段人类DNA的载体系统是有益处的。酵母人工染色体(YAC,正如第五章所讲)可克隆从100 000~1 000 000bp的DNA片段,这已成为物理图的标准“货币”(standard currency),尽管还有些技术问题。然而,将来可能会有更好的载体系统。因此,可以运用在未来不会过时的,同时也是任何一个有兴趣的研究者马上可以得到的“试剂”(载体)构建物理图。所以,序列标签位点(sequence-tagged site,STS)就成为现代物理作图的常用标记。乍一看来,这似乎是一种奇特的、确定物理图的方法,但经过思考后就会发现该系统的优点很明显。

图10-4说明STS的概念。简单地说,它是基因组中独特的位点,可用两个引物鉴定,该引物可用于聚合酶链反应,扩增人类DNA的一个或仅有的一个位置。引物的序列不会由于载体技术的进步而改变(因为序列本身是不会变化的),所以它在特定的位点规定为界标。而且,这样的序列会立刻被贮存于公共电子数据库中,任何有兴趣获得该区域DNA的研究者可通过合成适当的寡核苷酸就能够简单地从这个区域获得DNA。

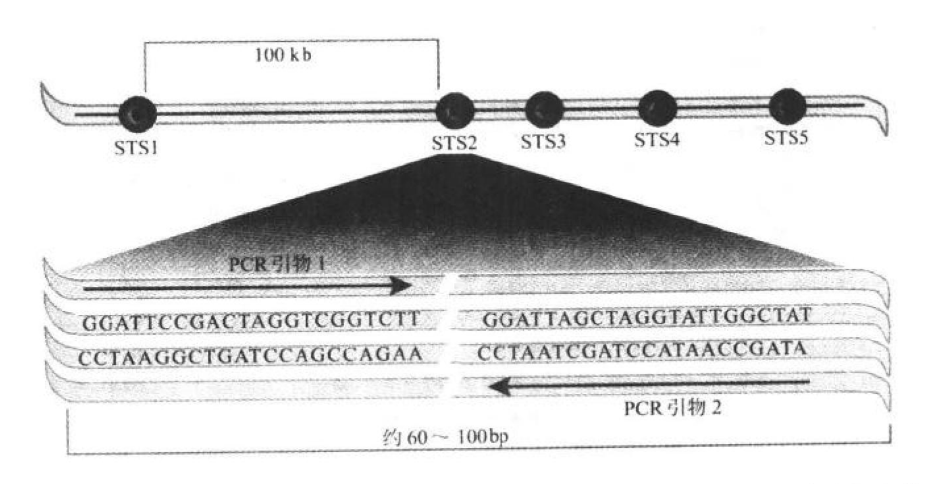


图10-4 序列标签位点(STS)是人类基因组独特的位点,可由两个寡核苷酸引物来确定。这些引物能用于聚合酶链反应来扩增特定区域的DNA。尽管载体系统在未来会发生变化,但确定STS的序列是基因组的一个基本特征,因此,这个“货币”对物理作图来说应是十分持久的。

这些STS可用在任何感兴趣的区域构建系列的重叠克隆,只要它们的间距很紧密。为更明确起见,这些STS的间距需要比在图中聚集的插入克隆片段的平均大小更紧靠在一起若干倍。图10-5表示在实践中这是如何进行的,此例是囊性纤维化基因的区域,一套32个STS被用来筛查整个人类基因组的YAC文库,而每个STS已检测若干阳性克隆。通过计算每个STS的每个克隆并记录其阳性和阴性的模式,就可重建STS标记的顺序和这套重叠YAC的顺序。在未来,如果可以得到另外一些大的插入克隆文库的话,这些STS同样能够用于从那些克隆中容易地构建一个图。

人类基因组计划现在的目标是构建一个由平均间隔100 000bp的STS所组成的人类

基因组的物理图，为的是能够产生像图 10-5 所示囊性纤维化基因（*CF*）区域那样的、所有染色体全部区域的图。事实上，这一工作进展迅速，在写此书之前就已得到 30 000 多个 STS，基因组的 98%以上被确定在重叠有序的 YAC 中。这使得对某一区域感兴趣的研究者会立刻从这个区域得到或重建克隆 DNA，从而节省数月乃至数年的工作。

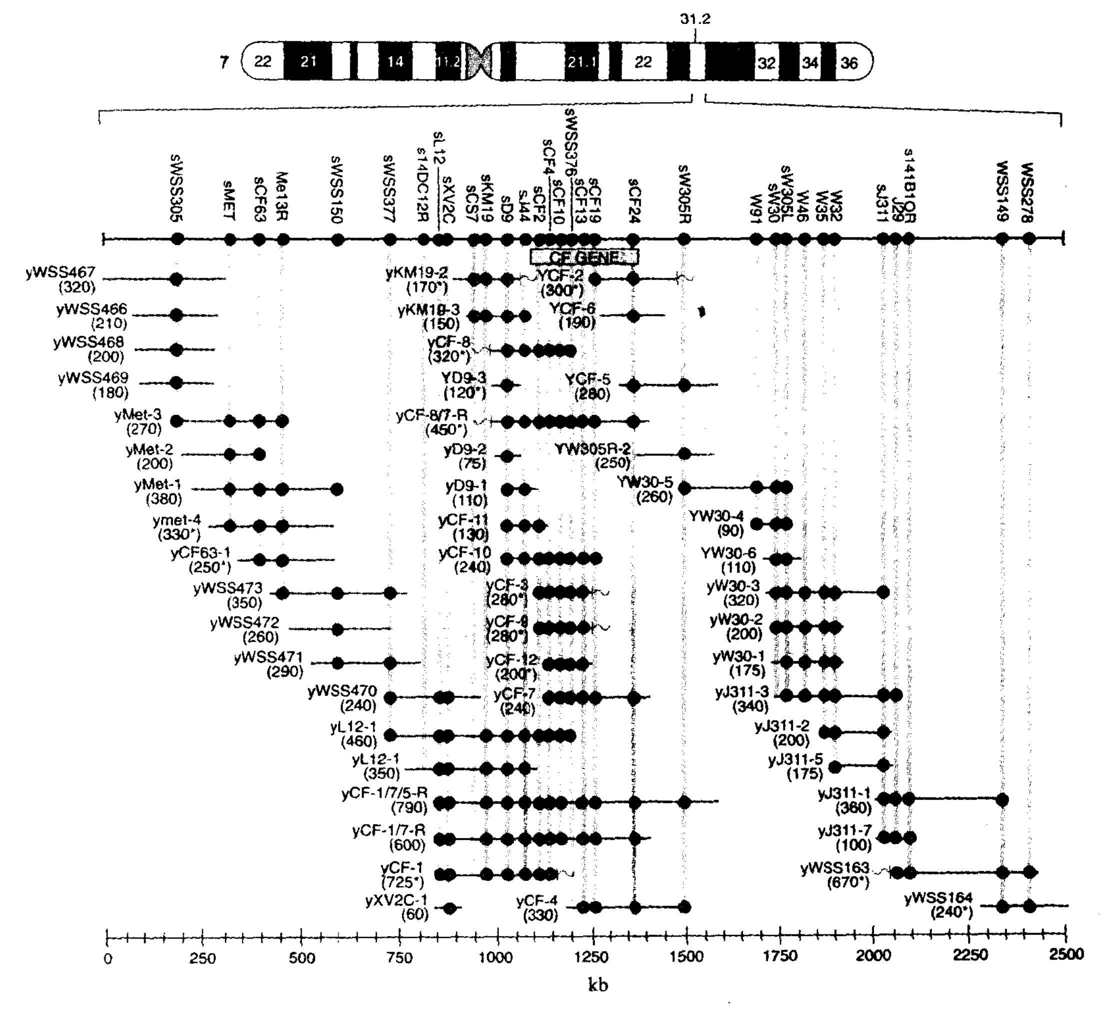


图 10-5　运用 STS 策略构建一套重叠克隆的例子。在此例中，该区包括囊性纤维化基因。STS（垂直虚线）用来筛查一个 YAC 文库，而每个 STS 则检测几个 YAC（水平黑线）。每个 YAC 的大小以 kb 在括号中标明，而其长度范围在 60～790kb。"镶嵌"（chimeric）YAC（即在 YAC 构建过程中，它们含有一个以上的、会偶尔连接在一起的人的 DNA 片段）的大小用 * 标明，而 ~ 表示非 7 号染色体物质。这套 YAC 是位于 2.5 Mb 区域内一套完全的重叠克隆（称为"重叠克隆群"）。

（三）测序

人类基因组计划的最终目的是要获得人类 DNA 全部 30 亿 bp 序列。当此计划开始之时，曾考虑 DNA 测序技术是否会迅速发展到足以产生如此大量的数据。然而，由于充

分注意自动化,优化所有的步骤以及规模微型化,现在显然可以确认这种努力是成功的。图 10-2 表明进入公共数据库的序列数目按指数率增长。

还有一个重要问题:即一旦掌握了全部的人的序列,是否有可能预示基因的位置呢?人们公认不到 5%的基因组 DNA 涉及到编码蛋白质,重要的是有可用的计算机演算系统,它能从图籍未载的大量 DNA 中识别编码信息(外显子)。外显子具有一定的分辨特征(第五章):它们的前面有一个拼接受体,通常含有一个可读框以及由一个拼接供体终结。然而,这些特征并不足以通过简单可见的检查从大量的 DNA 中找出编码区。近来精尖的神经网络方法已发展了计算机演算系统,该系统已成为在早先曾未研究过的 DNA 区域鉴定 100bp 或更大的外显子的有力工具。图 10-6 显示了称为 GRAIL 的计算机演算系统所输出的、关于 16 号染色体 17 000bp DNA 序列的一个例子。该区含有多囊肾病的部分基因,由计算机预测的外显子位置和这个外显子的实际位置非常一致。

考虑到外显子具有特殊价值,cDNA 测序引起人们极大的兴趣。原先在私人部门而近来在公共领域的一些研究组已经确定了成千上万个 cDNA 的部分序列,现在正在努力运用每一 cDNA 的 3′非翻译区(通常称为 EST, expressed sequence tags, 或表达序列标签)作为上述物理作图策略的 STS,将所有这些序列(cDNA)定位于物理图的正确位置。cDNA 和基因组测序的信息相结合将会很强大有力。曾有某些争议认为既然 cDNA 含有编码区,基因组序列则太多。然而,它假设能够鉴定人的全部 cDNA,但对于含量低或只在那些研究难以得到的组织中表达的 cDNA 来说,可能是困难的。而且,cDNA 测序将不提供决定哪些基因在哪些组织中表达的调节序列的重要信息。

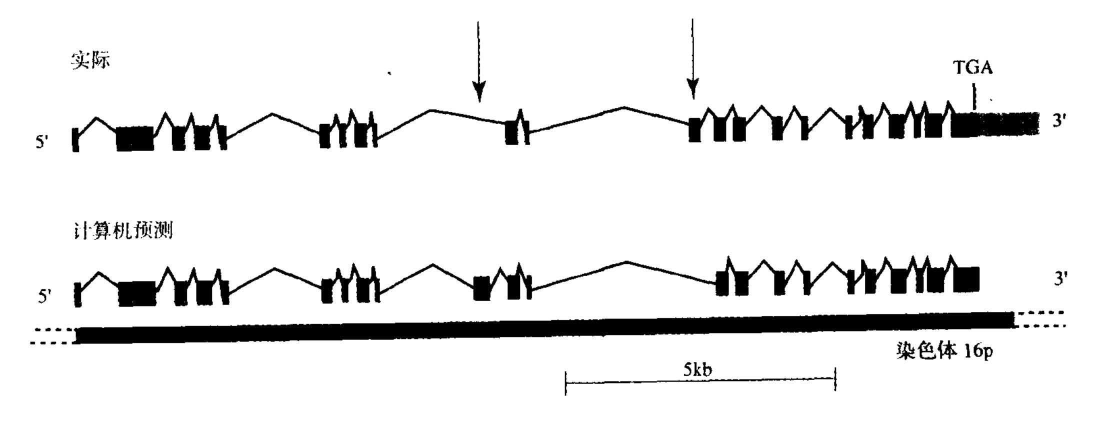


图 10-6 计算机演算系统现可用于扫描大区域的基因组序列并以合理的精确性预测编码区(外显子)的位置。在此例中,对 16 号染色体的 17kb 区域进行了分析,该区域含有常染色体显性的多囊肾病的部分基因。基因的实际结构如上所示,是由 cDNA 分析推断得出的。实体盒代表外显子,而连接它们的细线是内含子序列。终止密码子(TGA)出现在最后一个外显子。下行是计算机通过 GRAIL 演算系统仅仅根据基因组序列分析预测的编码区位置。显然,虽然计算机预测了一个并非真实的外显子(左箭头)和丢失一个外显子(右箭头),但其一致性十分好。

(四) 模式生物体

尽管完整的人类基因组序列是很有用的,但也仍存在许多难题。与其他更适合研究

分析的模式生物体进行比较将有益于序列的解析及其实践应用。因此,人类基因组计划从一开始就支持对某些其他生物体的制图和测序,如表 10-1 所示。仔细地选择了这些生物体是基于其大量的遗传信息和表型信息的可用性。除小鼠的基因组与人的基因组大小相同,因而全部以竞争完成之外,其他模式生物体基因组的大小比较容易操作。它们的基因密度高于人的基因密度,所以从其序列中可获取更多的信息,人类基因组计划的头几年所做的许多测序工作促进了这些模式生物体的基因组研究。酵母和大肠杆菌(*E. coli*)的基因组序列现已完成,而线虫(*C. elegans*)的基因组序列有望在 1998 年完成。

表 10-1 人类基因组计划研究的生物体模式

生 物 体	基因组大小	估计基因数目	基因/Mb
小鼠	3.0×10^{9} bp	80 000	27
Drosophila(果蝇)	1.2×10^{8} bp	10 000	83
C. elegans(线虫)	1.0×10^{8} bp	13 000	130
酵母	1.2×10^{7} bp	5 800	483
E. coli(细菌)	4.6×10^{6} bp	4 288	932

(五) 公共数据库

由人类基因组计划得到的制图和序列的信息正以惊人的速度积累,它只有在公共数据库中被科学团体免费得到才是有用的。为适应这一需要,已设计并实行了高精尖的电子版。同时美国政府支持的人类基因组计划的研究人员需要将他们的实验室数据在 6 个月内储存到这些数据库中,而不管这些数据是否已在杂志上发表。本章的附录提供了有用的电子信箱,可获得这些有价值的信息,现在任何人利用国际互联网均可免费利用。

三、伦理的、法律的和社会的意义

人类基因组计划的目的是要揭示我们人种遗传结构的大量信息。这将相当大地加快揭示个体疾病易感性致病基因的速度,这对人类健康具有很大的潜在利益,但也可能会滥用。虽然人类基因组计划实际上并不存在任何新的遗传困境,但是最终加速发现的步伐是令人满意的,即计划的设计者认为这项研究对伦理、法律和社会的影响的最大保证是正当的。这一努力是人类历史在伦理上最大的研究投资,即一般称之为伦理的、法律的和社会的影响项目(Ethical, Legal, and Social Implications, ELSI)。

在人类基因组计划的头几年已鉴定了一些最优先的 ELSI 问题,其目的是为了研究、审议,如果需要的话,还有立法的动议权(legislative initiatives)。包括:

1.公正的运用遗传信息。遗传信息对个体有相当大的价值,就像一个人了解他或她对某种早期诊断即可治愈的疾病(例如结肠癌)具有患病风险一样。但是这些信息也可用来歧视一个人。例如:允许保险公司中断与某些"早先存在疾病"(preexisting condition)保险者的保险项目的健康保险系统通过否定某些根据遗传测试发现在未来患病风险增

高的个体的保险项目表现出明显的歧视。这样的个体可能是完全健康的,而且事实上从医疗随访和预防保健中可以最大获益。因为我们之中无人有机会选择来自我们双亲的基因,所以利用这种信息来确定我们的健康保险项目在遗传上是不公正的。在写这本书的时候,由于 ELSI 项目的促进,正在努力由州和/或联邦立法通过来纠正这种可能的歧视。1996 年通过的健康保险实施和义务法令(Health Insurance Portability and Accountability Act, HIP & AA)为团体健康保险市场的个人提供了重要的保护措施,但仍有漏洞需要堵塞。堵塞漏洞的紧迫性随着每个新基因的发现而增长。毕竟我们所有携带的"早先存在疾病"将在未来被认识,而这种歧视损害的可能性将愈加深远。

同样地,在就业时潜在的遗传歧视已引起了重要的关注。在这方面也得到一些进展:1995 年,就业机会均等委员会(Equal Employment Opportunities Commission, EEOC)根据 ADA(Americans with Disabilities Act)组织发布了一项正式规定,即禁止这种歧视。但联邦立法仍需提供更好的保护。

2.隐私。一个相关的问题是关于个体的遗传信息未经本人同意能够传递给他人的程度。许多调查研究已经表明大多数个体对这类信息是敏感的,而且没有特定的许可,这些信息是不能利用的。然而,世界上大部分地区的医学记录系统并非都是安全的,而个人的遗传信息未经患者同意很容易被第三方获得。故这里再次提出,立法动议权正在发展一个能够更好地保护个人敏感性医学信息隐私的系统。

3.遗传检测的安全性和有效性。遗传学发现的数目日益增多,其中每一发现都开辟了预示个体及其后代未来患病的诊断检测的可能性,关注实验室检测的质量控制以及为确定新检测方法适于当前医学的时间所采取的手段,正与日俱增。在美国一些团体对这些问题非常感兴趣。消费者希望的是确保这种检测是安全和精确的,而且一旦信息可能有价值的话,希望可以及时加以利用。一些专业组织如美国医学遗传学学院(The American College of Medical Genetics)、美国人类遗传学会(The American Society of Human Genetics)和国家遗传咨询者学会(National Society of Genetic Counselors)等在了解遗传学发现引起最小风险和最大利益方面,都有很大的投入。美国的联邦机构如食品药物管理局(Food and Drug Administration)和卫生保健财政局(Health Care Financing Administration)负责诊断检测的调节义务,遗传检测比实验室其他形式的试验更加需要认真地监视。总之,一个个体只需测试一次,使其对错误的忍受几乎为零。测试的结果不仅对患者而且对其亲属都有影响。此信息可以是高技术的,同时必定与正确的遗传咨询相结合而获得理想的利益。正由于这一切,人们以相同的角度考虑新的遗传检测和新药是适当的。实验室合成的新药物在普遍投放医疗实践之前,一定要通过仔细的临床试验来获得其有效性和毒性的信息。同样地,许多人认为一个新的遗传检测在以一种不受约束的方式使用之前,应对其研究基础进行评价。

这些题目只是 ELSI 项目前三位的重点范畴。许多其他的争议还在前头,特别是在决定这一技术正确用于产前诊断时。那些困境中的一些问题将在第十四章中叙述。

四、对未来医学的重要性

人类基因组计划的科学目标被认为是遗传学整个领域的"建设性基础结构",同时也

已经加速了基因发现的步伐。图10-7所示为分子医学发现的进程,这些发现的大量出现正是这种努力的结果。由人类基因组计划所提供的图和序列的信息和特定疾病高频发生家系DNA的收集相结合,应用定位克隆的原理,应该能够鉴定几乎全部可遗传表型的易感基因。这些基因的发现将很快提高症状前检测的可能性。在具有高外显率的单个孟德尔式基因存在的情况下,这种检测可预示一个个体及其后代具有未来患病的高风险。然而,在多基因病或外显率相对低的疾病中,检测将不能提供是与否的回答,而是增高或降低的相对风险。对于许多疾病来说,有关风险的信息是有用的,因为它可设计一个医疗监视计划和降低这种风险的生活方式计划。例如,某人携带有如第十一章描述的一个直肠癌错配修复基因的生殖细胞突变,他将从40岁开始的、连续每年进行的直肠镜检查中充分受益,所以在进展为致命性侵袭性肿瘤许多年之前检查并切除结肠息肉。

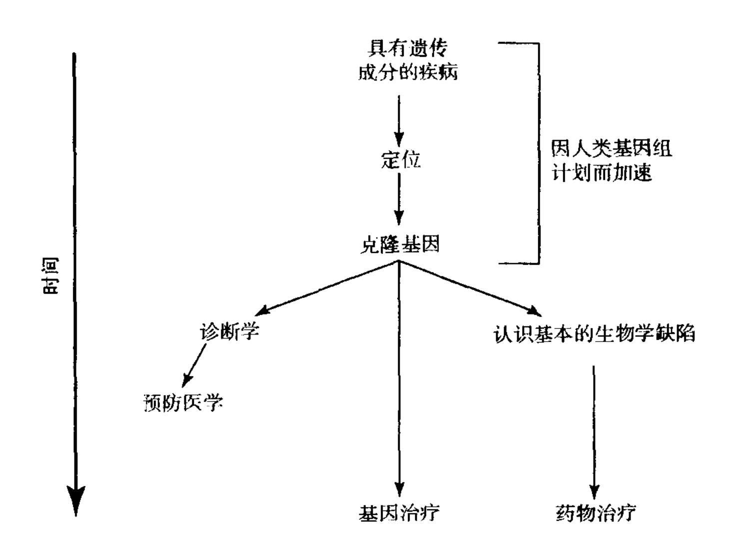


图10-7　尽管时间的间隔可能是几年甚至几十年,但在分子水平认识人类疾病的进展遵循一种可预示的模式,如此图从上到下所示。人类基因组计划的诞生极大地加快了运用定位克隆策略鉴定人类疾病基因的能力。鉴定一个基因后,对于发现高风险的那些个体可以设计预防医学策略。然而,从长远来看,基因发现的最大希望将是新的而且更有效的治疗方法的发展。

然而,并非所有的遗传病都与这种有用的预防医学策略相关。Huntington病就是一个典型的例子,现在能获得该病的精确的诊断信息,但尚无可以利用的干预措施来延迟或避免携有动态三联体重复的个体最终出现症状。许多其他的疾病也属此类。值得怀疑的是,是否大多数个体想要知道他们有关这些疾病的未来风险的信息。

但重要的是,基因发现的根本目的并非发展更好的诊断信息。正如图10-7所示,人类基因组的长期希望是这些基因的发现将导致更好的治疗。在某些例子中,利用基因本身治疗疾病是可以实现的(基因治疗,第十三章)。在另外一些例子中,基因的发现可提供充足的、关于基本的生物学和细胞缺陷的信息,使人们可以开发比现在通用的大量的经验药物方法更为成功而合理的药物治疗。但是在图10-7中值得注意的是:基因发现和治疗性进展之间是难以预测的长期时间。在某些例子,这一间期可能是几年,而对于另

一些疾病而言,在深刻认识该病精确的遗传学和分子生物学基础并开发有效的治疗方法之前,可能已经过去了几十年。在有力的诊断方法和最终的成功治疗之间的时间之窗在未来的几十年中将给医学实践带来许多挑战。然而,以最富生命力的途径向前发展可能是未来最大的希望,否则将会推迟最终认识和医治令我们痛苦的重大疾病。

推 荐 读 物

Bishop JE, Waldholz M. Genome. New York: Simon & Schuster, 1990. *An engaging and accessible description of the steps leading up to initiation of the Human Genome Project by two reporters for the Wall Street Journal.*

Cook-Deegan R. The Gene Wars. New York: Norton, 1994. *An autboritative firstband account of the scientific and political roots of the Genome Project.*

Collins FS. Ahead of schedule and under budget: The Genome Project passes its fifth birthday. Proc Natl Acad Sci USA 1995; 92: 10821 ~ 10823. *An update (as of November 1995) of progress of the project.*

Hudson KL et al. Genetic discrimination and health insurance: An urgent need for reform. Science 1995; 270: 391 ~ 393. *A call for specific action to prevent the use of genetic information to deny health insurance coverage or set exorbitant premiums.*

Collins FS. Sequencing the human genome. Hospital Practice 1997; 32: 35 ~ 53. *An update on progress and application to the future of medicine.*

附录:基因组信息数据库

DNA Data Bank of Japan
http://www.ddbj.nig.ac.jp/

European Bioinformatics Institute (EBI)
http://www.ebi.ac.uk/

Genome Data Base
http://gdbwww.gdb.org/

National Center for Biotechnology Information (NCBI)
http://www.ncbi.nlm.nih.gov/

Mouse Genome Database
http://www.informatics.jax.org/

Online Mendelian Inheritance in Man (OMIM)
http://www.ncbi.nlm.nih.gov/Omim/

Saccbaromyces (yeast) Genome Database
http://genome-www.stanford.edu

Généthon (French genome center)
http://www.genethon.fr

US Department of Energy (DOE) Human Genome Program
http://www.er.doe.gov/production/oher/hug-top.html

National Human Genome Research Institute
http://www.nhgri.nih.gov/

第十一章 癌遗传学

另一种可能性是，每一个正常细胞中有一种特殊的排列可以抑制细胞分裂，只有当某种特殊刺激解除这种抑制后，细胞才能分裂。假定存在一些抑制分裂的确定的染色体，就和我的基本观点极其一致……如果那些"抑制性染色体"丢失将引起肿瘤细胞的无限生长……另一方面，假定还存在促进分裂的染色体，也可令人相信这个假设。据此假设，这些染色质部分的作用，通常很弱，当受到某种刺激加强时，细胞就发生分裂；由此可以推断恶性肿瘤细胞的快速无限增殖的趋势，是由于促进分裂的染色体的持久优势所致。

——Theodor Boveri，1911

上述引文，初看起来写得似乎很极端而且难理解，实际上它是对人类癌遗传学的神秘而确切的阐述，这是 Boveri 在证实或否定其观点的有用手段问世 60 年前提出的。在人类分子遗传学的所有领域中，可以说没有任何领域能比肿瘤发生机制的研究更令人兴奋。这种机制仍未完全了解，并将在今后的几十年中继续深入。然而，由于它在人类疾病中基本的重要性，所以用一整章的篇幅来阐述此问题似乎是适当的。努力理解癌症发生的分子基础也有助于对前几章中提出的基本原则的认识，因为它将孟德尔式遗传、细胞生物学、细胞遗传学、病毒学和分子生物学集合于一体。

一、大多数人类癌症受遗传上的影响，但不呈孟德尔式遗传

由于营养不良和感染性疾病的灾祸在世界许多地区得到较好控制，于是癌症便成为健康的主要问题。随着心脏疾病死亡率的逐渐降低，癌症将很快成为发达国家中死亡的首要原因。尽管在少数特殊类型癌症的治疗方面获得显著的成功，但在过去的 30 年中，总的癌症死亡人数仍然令人失望地保持在高水平（虽然近来报道显示这些数字可能开始下降）。当然癌症不是单一的疾病，而是一组不同的疾病，其发生的部位、组织学类型以及恶性程度均有自己的特点。

大多数人类癌症不能用单基因遗传方式来解释。然而，这并不是说常见的癌症没有遗传因素。数十年的研究表明，某种特殊类型癌症患者的一级亲属发生同类肿瘤的风险通常略有增高。这类分析由于环境因素在许多癌症中所起作用而复杂化（例如，吸烟和肺癌）。多因素模式（第四章）认为环境和一些基因与风险有关，但要真正了解常见癌症遗传易感性的机制尚有待进一步认识最近发现的易感基因的特性，鉴定其他的相关基因，并加深对这些基因和环境间相互作用的理解。

二、癌症是一种遗传病

尽管任何一个患者发生癌症通常不能归因于单一遗传基因的异常，但在细胞水平癌症基本上是一种遗传病。每一个人类细胞携带非常复杂且高度调节的遗传程序——调控正常细胞的生长和分化。癌症通常起因于这个正常调节方式被破坏，导致细胞无限制地生长和增殖，认为就是恶性肿瘤。通过局部生长蔓延并转移到远隔部位，癌破坏正常器官功能引起严重的疾病并且通常导致死亡。正如我们将要看到的那样，癌症的充分发展通常需要在一个特定的体细胞内多个、连续的遗传突变的积累。这些体细胞突变仅限于肿瘤，而不存在于此个体的正常细胞。相反，生殖系突变则存在于某一个体的所有细胞中，并且可以传递给后代。实际上，一些在癌细胞中通常为体细胞突变的基因，在生殖系中也存在突变，并产生遗传特异的孟德尔式癌症易感综合征。

在过去的 20 年中，癌症研究者在鉴定许多基因方面取得了显著的进展，这些基因突变时能够参与癌细胞的发展。从这些大量的工作中，癌症发病机制的概貌正开始显现。

三、癌的克隆特性

在细胞水平癌症遗传基础的最强有力的证据之一，是观察到几乎所有的癌症都是单克隆起源的(通常缩写为“克隆的”，clonal)；即表明某一患者的所有癌细胞都起源于单一的原始前体细胞。据推测，一个关键性的突变事件(或一系列事件)导致单个细胞转化为癌细胞，其随后的失控增殖产生大量的子细胞最终形成肿瘤。癌克隆性的证据有多种类型。对白血病或由恶性淋巴细胞组成的淋巴瘤的分子水平分析证明，所有的肿瘤细胞都携带同样的免疫球蛋白基因重排或 T 细胞受体基因重排，这提示它们起源于单一的原始 B 或 T 细胞，如图 11-1 所示。应用识别有活性或失活 X 染色体的各种方法对源于女性癌症的研究表明，在一个体肿瘤的所有细胞中几乎总是同一 X 染色体失活，这表明它们起源于一单个细胞(见第八章)。最后，在肿瘤中已发现了许多不同的细胞遗传学异常。在肿瘤的所有单个细胞中通常可见相同的标记染色体，这再次证明了常见的克隆起源于单一、原始的恶性细胞。费城染色体的例子如图 11-2 所示，该染色体可见于 90% 以上慢性粒细胞白血病(chronic myelogenous leukemia，CML)患者的癌细胞中。

四、显性转化基因：“癌基因”

大多数肿瘤的克隆性表明遗传的突变可以引起一种恶性表型。鉴定与这种现象有关的基因一直是研究的热点课题。人们可以假设除非它们发挥重要的正常功能，否则进化将不会允许这样的危险基因存在，并且还可进一步假设这些功能与控制细胞生长有关。已证实这两种推测都是正确的。这个难题的一些线索已由许多未曾预料到的研究领域提供，这需要我们简单地回顾一下 RNA 肿瘤病毒。虽然这类病毒是人类癌症的罕见原因，但对其他有机体中生物学特性的研究却为理解癌遗传学提供了重要线索。

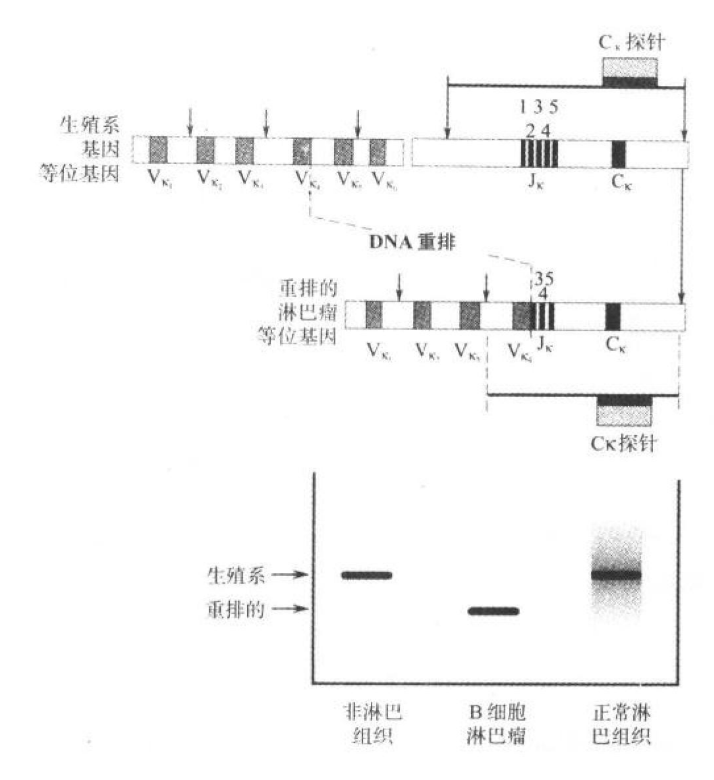


图 11-1　B 细胞淋巴瘤的克隆性。图上方的是 κ 轻链基因的示意图解。此基因在生殖系中未发生重排，但在 B 细胞中发生 DNA 重排并产生一个可表达的基因序列。每一个 B 细胞经历了本质上独特的重排，因而必然产生抗体多样性。垂直箭头示 *Bam*HⅠ酶限制性位点的部位。位于下方的 DNA 印迹图说明非淋巴组织（泳道 1）、克隆的 B 细胞淋巴瘤（泳道 2）或非克隆的正常淋巴组织（泳道 3）经 *Bam*HⅠ消化用 Cκ 探针杂交后的结果，Cκ 来源于 κ 轻链的恒定区域。非淋巴组织产生单一条带，其长度由位于 Cκ 任一侧的 *Bam*HⅠ位点所确定。然而在源于单一的原始 B 细胞的淋巴细胞瘤中，那种细胞特异的信号将存在于该肿瘤的所有 DNA 中，因为 κ 基因的某一特殊重排存在于这一原始转化细胞中。因此出现了不同于生殖系中的带。其确切大小取决于哪一 Vκ 区域参与细胞中与恶变有关的特殊重排。在这种情况下，发现一重排带是克隆性的指征。另一方面，第 3 泳道显示对正常淋巴组织的分析结果，它由大量不同的重排 κ 基因和一些未重排的等位基因组成。未出现明显的重排带，但却见到由许多不同大小的重排等位基因组成的背景性连续的带谱，并伴有残留的未重排等位基因形成的微弱的生殖系条带。

（一）转化性反转录病毒

以 Peyton Rous 对鸡的先驱性研究作为开端，在鸟类和哺乳类中发现了许多能诱发癌症的可传染的因子。这些研究最终证明许多可传染的因子可以归为一类称为**反转录病毒**（retroviruses）的 RNA 病毒。这类病毒的基因组由单链 RNA 组成。如图 11-3A 所示，一旦进入细胞，反转录病毒即脱去外壳，利用称为反转录酶的特殊酶，将 RNA 复制为 DNA。这种酶由反转录病毒携带并编码。然后 DNA 拷贝插入宿主基因组。在其中它可能保持

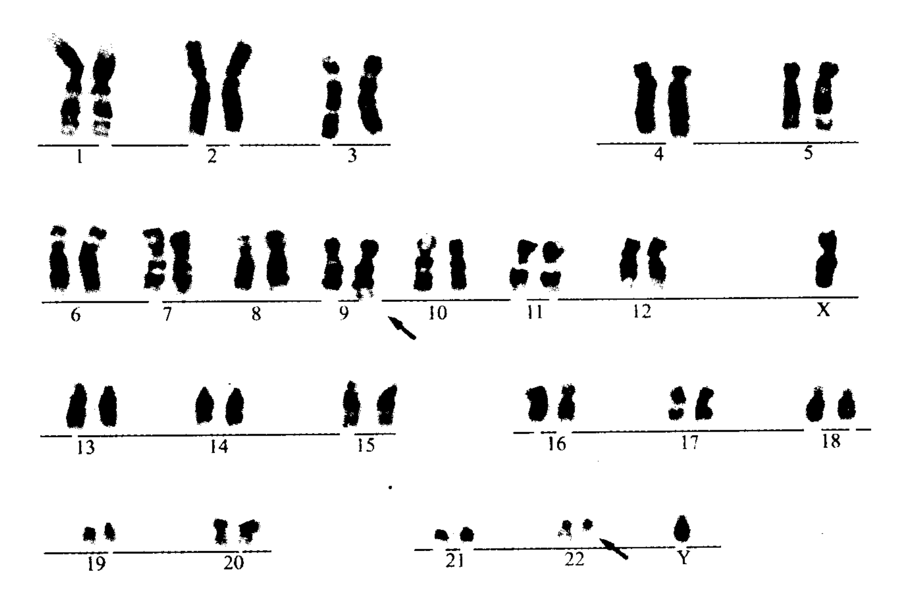


图 11-2 在慢性粒细胞白血病(CML)中所见显示典型的 t(9;22)易位核型。缩短的 22 号染色体称为费城(Ph^1)染色体(承蒙 Dr.T Glover 允许)。

静止或可能转录产生自身的多个 RNA 拷贝,然后再进行包装,导致进一步轮回的感染。常见的反转录病毒基因组的简图如图 11-3B 所示,表明存在三种基因:*gag* 基因编码病毒的核心蛋白,*pol* 基因编码反转录酶,*env* 基因编码被膜蛋白。包装序列(Ψ)是转录的 RNA 分子包装成病毒颗粒所必需的。

仅鉴定的少数人类反转录病毒包括 AIDS 病毒(人类免疫缺陷病毒,human immunodeficiency virus,或 HIV)。像 HIV 这样完整的反转录病毒通过促进细胞增殖和增加体细胞突变的机会,或插入基因组后引起邻近基因的激活引起免疫缺陷而间接引起肿瘤发生 。然而,另一类反转录病毒称为急性转化反转录病毒,通过病毒本身携带的转化基因的作用更迅速地产生肿瘤。这些基因替代一些或所有的内源性病毒基因,形成一种复制缺陷型病毒,需要一种完整病毒的辅助才能繁殖。一般用最初鉴定的病毒的三个字母缩写来命名这些转化基因。例如,猴肉瘤病毒转化基因称为 *sis*,鸟类髓细胞瘤病病毒转化基因叫做 *myc*,等等。从多种鸟类和哺乳类中已鉴定了许多急性转化反转录病毒,尽管至今尚未发现这些反转录病毒与人类肿瘤直接有关。

(二) 用转化实验克隆人类癌症基因

尽管反转录病毒与人类癌症的关系尚有质疑,但反转录病毒研究者仍在不断地鉴定感兴趣的 RNA 肿瘤病毒。在 20 世纪 70 年代中期,其他研究者开始应用分子生物学技术对人类癌症进行更直接的研究。如果细胞基因的特异突变能够引起人类癌症的恶性表型,并且如果那些突变在细胞水平发挥显性作用,那么将恶性细胞的 DNA 序列转移到正常细胞预期可导致其转化为恶性表型。1981 年这个实验在三个实验室同时获得成功并

A.

1.吸附并进入
2.脱壳
RNA
细胞膜
3.反转录
核
RNA
DNA
DNA
DNA
8.非溶解性出芽
4.整合
（分裂的细胞）
基因组 mRNA
5.转录
7.包装
mRNA
mRNA
mRNA
6.翻译
外壳蛋白
内部结构蛋白
反转录酶
病毒的 RNA 基因组

B.

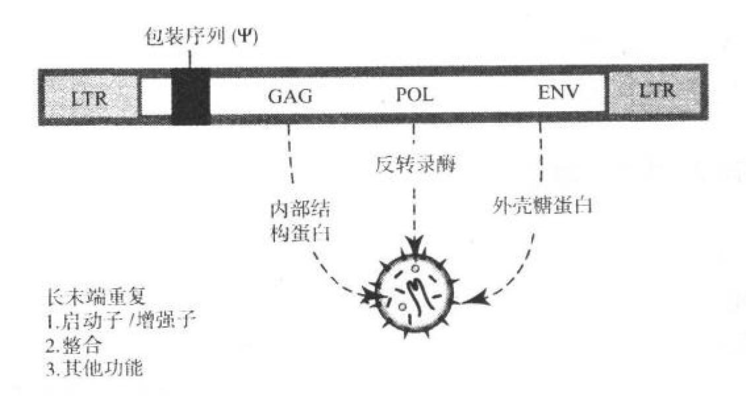


图 11-3　A. 反转录病毒的生命周期。B.典型的反转录病毒基因组图。

取得的重大进展。此实验的基本原理如图 11-4 所示。从人肿瘤细胞系制备 DNA;最初用于实验的是称为 EJ 的膀胱癌。其 DNA 被转染到称为 3T3 的小鼠成纤维细胞系中。

当在培养皿内生长时,像 3T3 这样的非恶性细胞彼此接触时将停止生长(称为“接触抑制”,contact inhibiting),而恶性细胞一般继续生长,它们层层堆积形成转化灶。当人 EJ 膀胱癌细胞系的 DNA 被导入小鼠 3T3 细胞并观察培养的细胞,出现了小鼠细胞的转化灶。而用非恶性来源的 DNA 时,则无这样的转化灶出现。而且,将来自转化灶的 3T3 细胞接种于免疫缺陷小鼠(称“裸”鼠,因为它们没有毛)时产生肿瘤结节,而非转化的 3T3

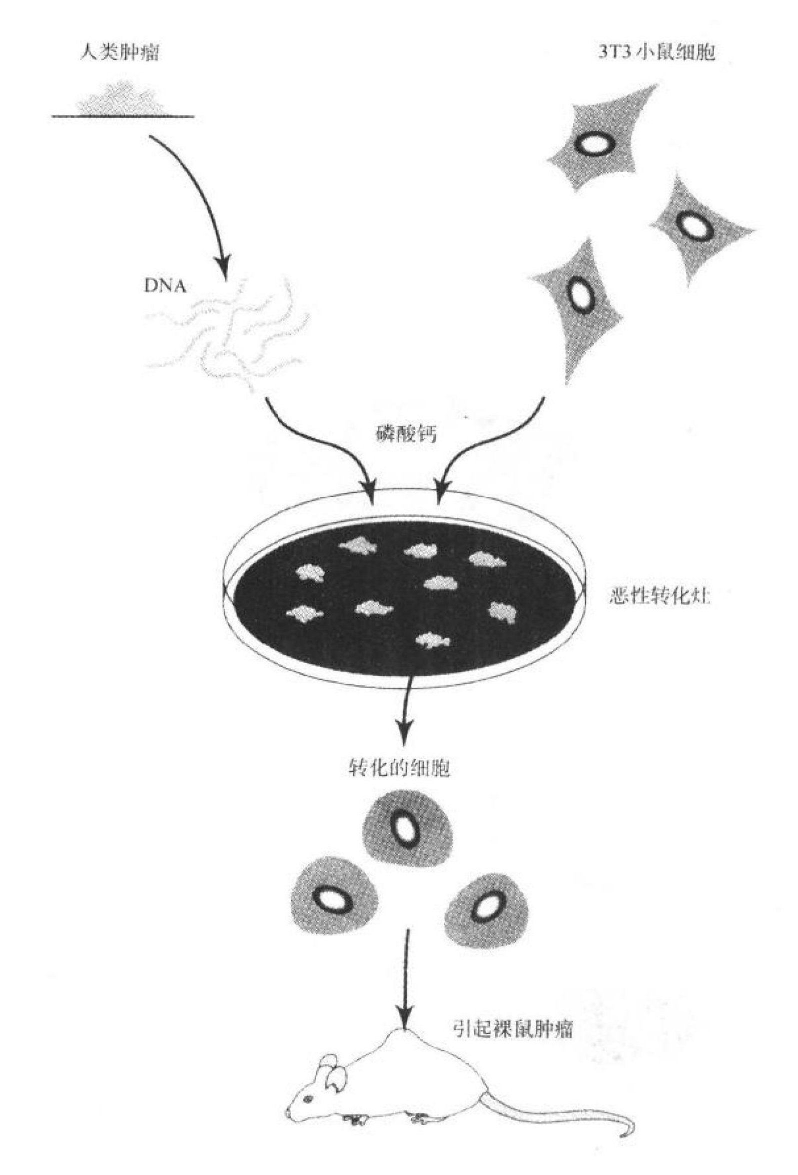


图 11-4 证明人类肿瘤 DNA 具有使小鼠 3T3 细胞发生恶性转化能力的转染实验。制备人类肿瘤 DNA,用磷酸钙法将其转染到 3T3 细胞。接受人类癌症基因的细胞发生转化并在培养皿上聚集,形成转化灶。当皮下接种时,这些转化的细胞在裸鼠体内也可形成肿瘤,而非转染的 3T3 细胞则不形成肿瘤。

细胞则不产生肿瘤结节。这些发现强有力地支持这样一个观点,即膀胱癌 DNA 确实含有显性作用的癌症基因,随即将其命名为**癌基因**(oncogene)。

为发现并从分子水平克隆这个癌基因,运用了如图 11-5 所示的巧妙策略。这个方法依赖于人类 DNA 中分散重复的 *Alu* 序列(第五章),它可与小鼠 DNA 相区分。从原代转

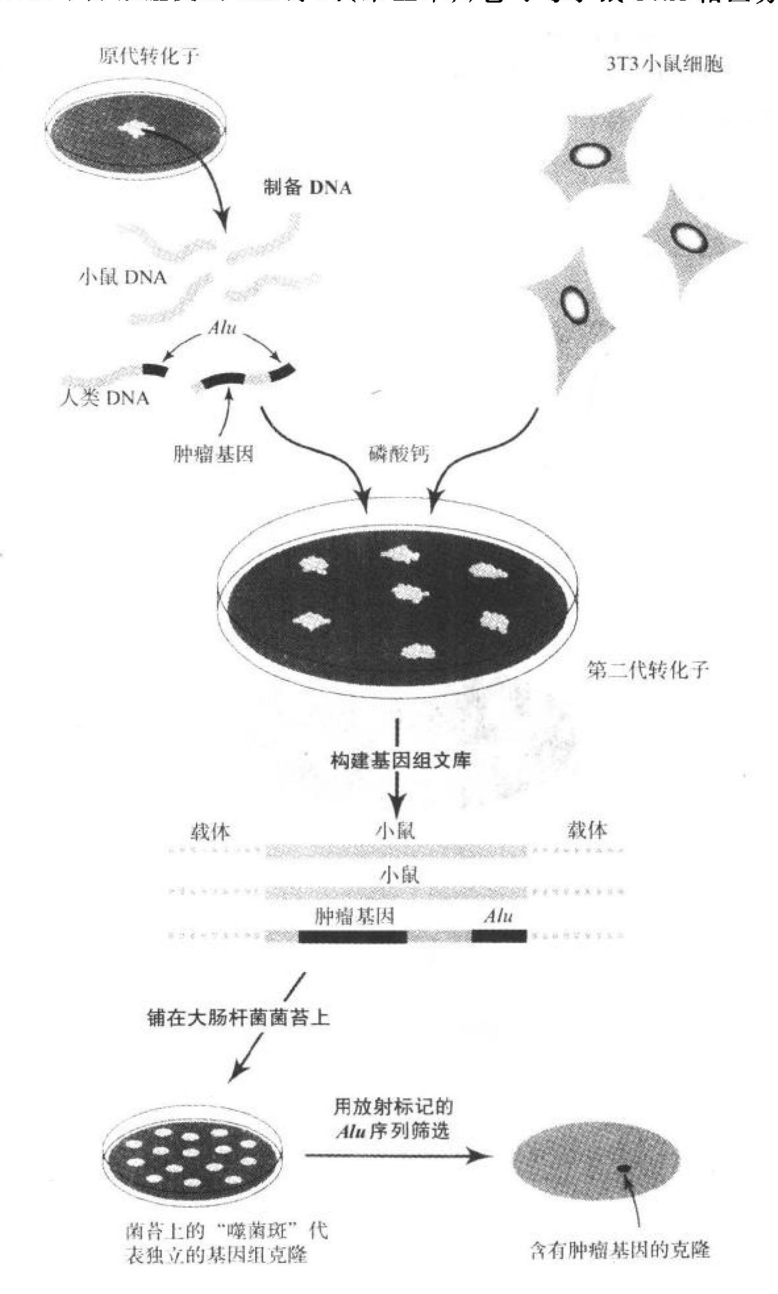


图 11-5 克隆显性癌基因。制备转化灶的 DNA(图 11-4),用这种 DNA 再转化小鼠 3T3 细胞("第二轮"转染,用于限制仍存在于转化的小鼠细胞中的人类 DNA 量),再制备所形成灶的 DNA,并克隆到 λ 基因组文库中,如第五章所述。然后用与人类特异的 *Alu* 重复元件相应的放射性标记序列筛查此文库来鉴定人的序列。

化灶制备的 DNA 含有许多人 DNA 的不同片段,用这些 DNA 对 3T3 细胞进行第二轮转化。此时,在第二代转化灶中的大多数 DNA 是来源小鼠的,只有一小部分含有癌基因的人 DNA 仍然存在。应用人类 *Alu* 序列作为探针,可将所有的第二代转化灶中共有的某一 DNA 片段进行分子克隆。将这种纯化的 DNA 重新转染小鼠 3T3 细胞时,形成转化灶的效率非常高,这表明已克隆了可靠的癌基因。当确定了这个转化基因序列后,这个研究组惊奇地发现,它与称为 H-*ras* 的反转录病毒转化基因序列极其相似。

那么,转化反转录病毒和肿瘤的癌基因这两个不同领域怎么会有关联呢?答案就在于如图 11-3A 所示的反转录病毒的生命周期中。当整合到宿主基因组后,能发生一种罕见的重组事件,结果反转录病毒被切开,并携带部分宿主 DNA 的片段。如果这个 DNA 片段恰好包含生长促进基因,如果反转录病毒通过应用其侧翼序列(称作 LTRs,见图 11-3A)中强有力的启动子"激活"了这一基因,则这种反转录病毒以后感染其他细胞将使其转化。这样的病毒将为它所感染的细胞提供一种生长优势,并因此而有所选择。另一方面,在人类肿瘤如 EJ 膀胱癌中发现的内源性癌基因,显然具有正常生长促进基因的激活功能,这是通过一种体细胞事件而与病毒无关,但结果却是相同的,即都发展为恶性表型。正常细胞中具有这种潜能的基因称为**原癌基因**(proto-oncogene)。

以后的实验表明,即使是从膀胱癌细胞系中克隆的激活的 *ras* 基因也不能转化所有类型的细胞。特别是对直接来源于活组织检查的人或啮齿类成纤维细胞进行培养时,在培养过程中发生"永生化"之前(如 3T3 细胞系那样),不能单独被这种基因转化。然而,两种癌基因相结合,比如 *ras* 加 *myc*,则可引起原代成纤维细胞转化。因此最初用于实验的小鼠 3T3 细胞不完全是正常的。这些观察与"恶性肿瘤的发生是一个涉及两个或更多阶段的事件"这一假说是一致的。以后我们将再次讨论这一主题。

(三) 细胞原癌基因及其功能

已鉴定了 40 多种原癌基因,最初是通过它们在急性转化反转录病毒中存在的活性形式(癌基因),或通过 DNA 转染实验从人类肿瘤中分离出来。如下叙述,其他一些基因则是通过 DNA 扩增或染色体易位使它们激活而相继被鉴定。原癌基因与促进生长有关的预言已经由一些特殊的例子得到证实。已证明 *sis* 癌基因与血小板生长因子(platelet-derived growth factor)亚单位的基因 *PDGF* 相同,PDGF 是一种由血小板和其他细胞产生的重要物质,它作为细胞生长和分裂强有力的刺激物,参与正常的创伤愈合。另一类原癌基因已证实是生长因子受体基因,该受体是存在于细胞膜上的结构,由一种特异的生长因子激活。生长因子与其受体结合引发一系列事件,导致有丝分裂信号传递至核内。*erbB1* 癌基因相当于表皮生长因子受体基因(the gene for epidermal growth factor recptor, *EGFR*),而 *fms* 癌基因与集落刺激因子 Ⅰ 受体(receptor for colony-stimulating factor 1)基因 CSFR1 相当,CSFR1 是一种促进骨髓中多种细胞分裂的生长因子。

其他的原癌基因如 *RAS*,可能是通过跨膜受体与核之间的信号传导通路而发挥其作用。最后,若干原癌基因包括 *MYC* 和 *FOS* 编码的蛋白质位于核内,可能直接参与 DNA 转录或复制的调控。这些原癌基因的分布情况总结于图 11-6。

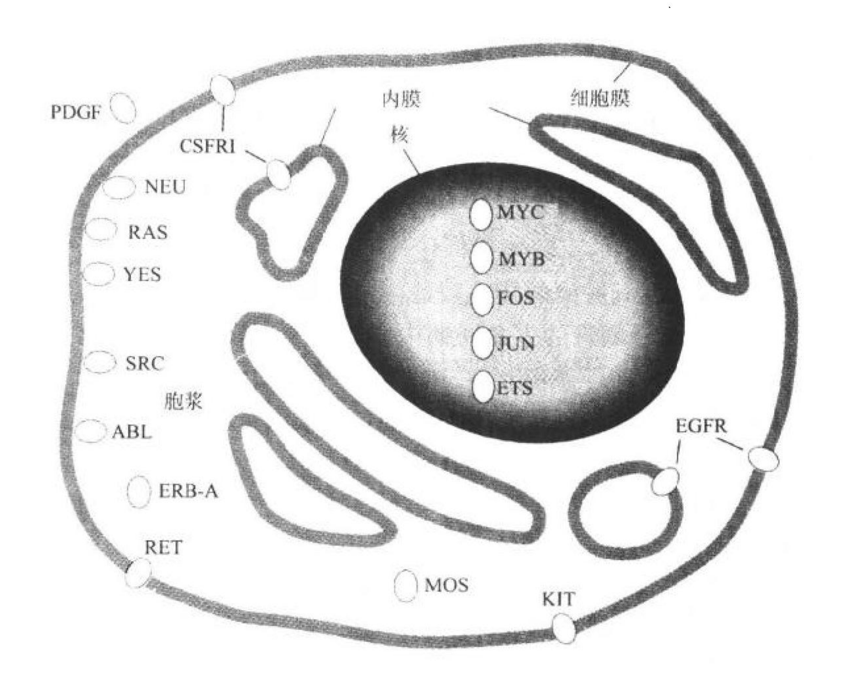


图 11-6 各种原癌基因蛋白产物在细胞内的定位。像血小板生长因子(PDGF)这样的一些因子是分泌性生长因子;其他一些,如表皮生长因子受体(EGFR),是位于细胞膜上的生长因子受体;另外一些,如 RAS,参与信号从膜向核的传导,还有一些如 FOS 和 MYC 则位于核本身。

(四) 癌基因是怎样激活的?

如果原癌基因是正常细胞生长所必需的序列,那么什么事件负责它们的激活及随后的失控? 可能有几种不同的机制。

1. 点突变

从人膀胱癌细胞中克隆的 *RAS* 基因最初是一个谜。当与非恶性细胞中其正常副本比较时,未发现该基因明显的重排,对该基因测序后发现了一个令人吃惊的微小异常,即该基因的第 12 密码子的点突变,将此位点的甘氨酸转变为缬氨酸。

另外,应用小鼠 3T3 细胞转化实验(图 11-4 和 11-5)对大量的肺、结肠肿瘤进行分析,克隆了激活的 *HRAS* 基因以及激活的 *KRAS* 和 *NRAS* 基因,它们是同一基因家族的有关成员。引人注目的是,所有这些激活的 *RAS* 基因的第 12、13 或 61 密码子含有点突变,表明这些位点对细胞增殖的正常控制一定是至关重要的。RAS 蛋白产物能以活性形式和失活形式存在,前者结合到 GTP 上,而后者可与 GDP 结合。显然,这些点突变阻止了 RAS 从活性形式向失活形式的转变,因此引起细胞生长失调(图 11-7)。

2. 扩增

原癌基因激活的另一种可能的机制是其蛋白产物显著的过表达。在某些情况下,这实际上是由于癌细胞中编码原癌基因的 DNA 序列扩增所致。DNA 扩增的片段一般是数百个 kb 长,通常包括位于癌基因附近的一些其他基因。这些扩增的 DNA 节段有时是可见的,在一条染色体上分散为独立的小片段,叫做"双微体"("double minutes")(图 11-8A)。当扩增的 DNA 插入一条染色体时,它则称为**均质染色区**(homogeneously staining region)或 HSR(图 11-8B)。虽然 HSR 和双微体在一些不同的癌中都可见到,但仅在少数病例鉴定了相关的已扩增的癌基因。在某些儿童较常见的一种神经母细胞瘤中 N-*MYC* 基因特征性的扩增就是一个例子。N-*MYC* 基因的扩增与预后不良有关。

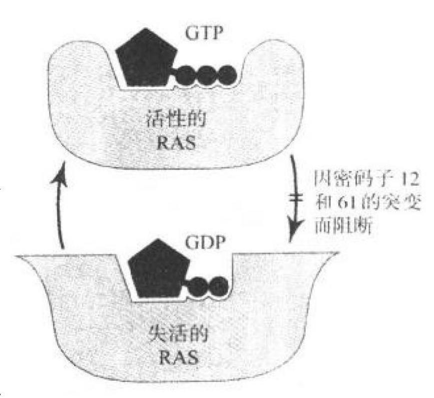


图 11-7　RAS 蛋白激活可能的调节机制。活性形式结合 GTP,而失活形式结合 GDP。显然第 12 和 61 密码子的点突变,使 *RAS* 转变为转化基因,阻止它从活性形式转变为失活形式,因此蛋白处于"锁定"位置。

3. 染色体易位

在某些显著的例子中,在原癌基因基因座或其附近发生的染色体易位可导致基因的激活。第一个这样的例子见于 Burkitt 淋巴瘤,该肿瘤是一种在美国罕见但在非洲某些地区常见的 B 细胞肿瘤。以往的细胞遗传学分析表明,大多数 Burkitt 淋巴瘤以 8 号和 14 号染色体平衡易位为特征。在 19 世纪 80 年代初期进行的基因定位发现一个有趣的问题,*MYC* 原癌基因定位于 8q24,而免疫球蛋白重链基因定位于 14q32,这些是与 Burkitt 淋巴瘤易位有关的区带(见图 11-9)。后来对许多 Burkitt 淋巴瘤进行分子水平分析证实,断裂点确实发生在 *MYC* 基因座或其附近和免疫球蛋白基因座内。结果使 *MYC* 基因座移至 B 细胞中转录活跃的基因即免疫球蛋白基因附近。虽然 *MYC* 基因激活的确切机制仍不清楚,但可能是通过这一事件以一种重要的方式使 *MYC* 基因调节紊乱,使得这种核癌基因失控,导致淋巴瘤转化。

另一个经染色体易位而使癌基因激活的例子是费城染色体,其核型如图 11-2 所示。这种特征性的染色体易位导致一种独特的嵌合基因产生,这种嵌合基因在 CML 的发病机制中可能起重要作用,如图 11-10 所示,这种易位将位于染色体 9q 的 *ABL* 原癌基因和位于染色体 22q 的 *BCR* 基因融合起来,由此产生的嵌合性 BCR-ABL 蛋白保留了 *ABL* 原癌基因的蛋白激酶活性,尽管其正常功能和表达形式已被与它融合的 *BCR* 基因所改变。这种已改变的 ABL 蛋白的异常表达可能是 CML 发生的一个重要原因。

对与特殊肿瘤类型有关的特异染色体易位的研究已鉴定了一些新的原癌基因及其激活的机制。一个癌基因和一特异肿瘤间的相关性也可能为某种癌症的起源提供重要线索。除了几乎所有的 CML 都具有的特征性费城染色体易位以外,已发现了若干其他特异相关性。15 号和 17 号染色体的独特易位将维甲酸受体基因与以前未知的称为 *PMR*

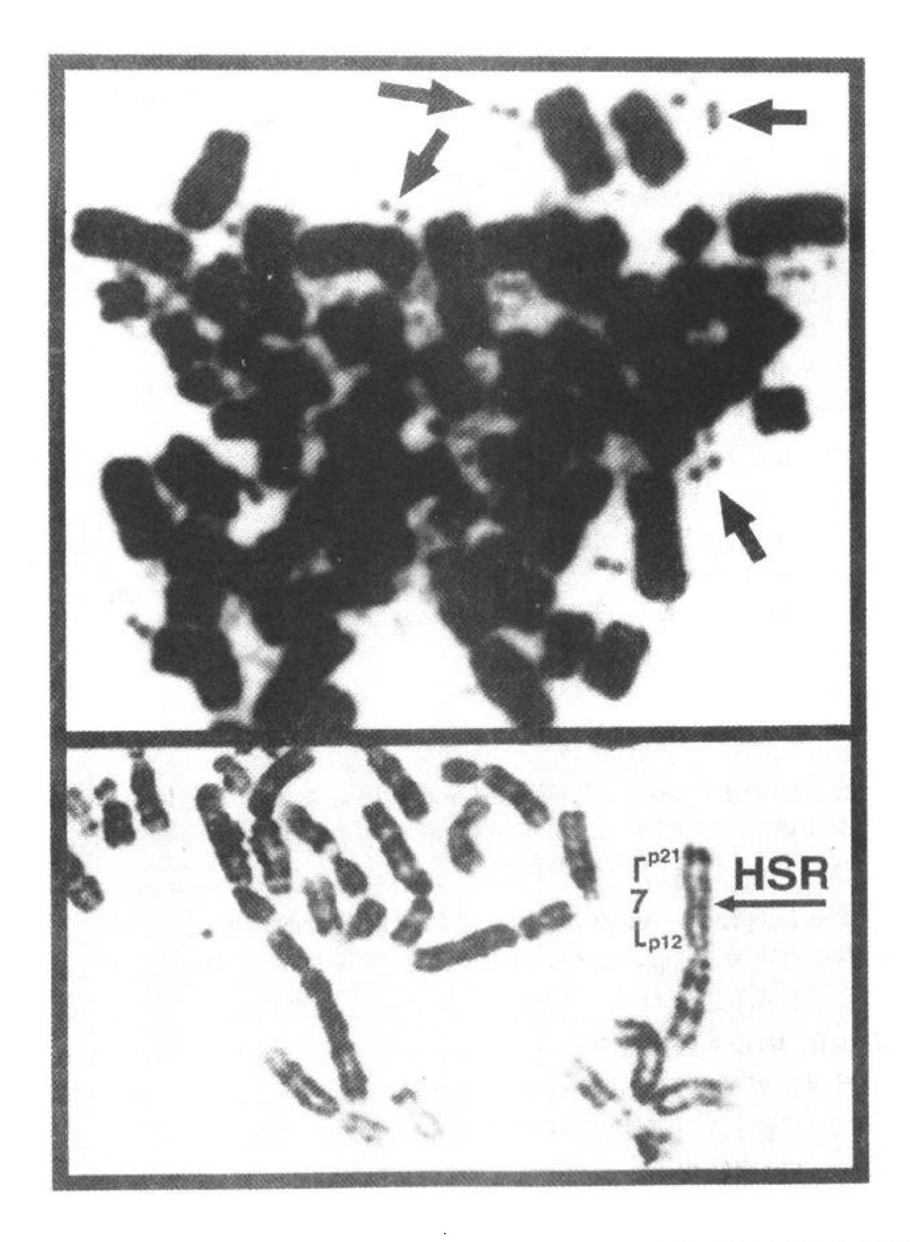


图 11-8　扩增激活原癌基因。A 用吉姆萨染色在此肿瘤细胞中期可见很多拷贝的"双微体"(箭头)。B 在这个 G 显带肿瘤细胞中期 7 号染色体短臂上可见 HSR(箭头)(引自 BeaudetAL, Scriver CR, Sly WS, Valle D, The metabolic and molecular bases of inherited disease. CD version . New York: Mcgraw-Hill, 1997.)。

的基因相融合，该易位几乎在所有的急性早幼粒细胞性白血病中都可见到。14 号和 18 号染色体间的特异易位是滤泡性淋巴瘤的特征，对此进行分析发现了 *BCL-2* 基因，后者是程序性细胞死亡(凋亡，apoptosis)调节中的一种重要蛋白。

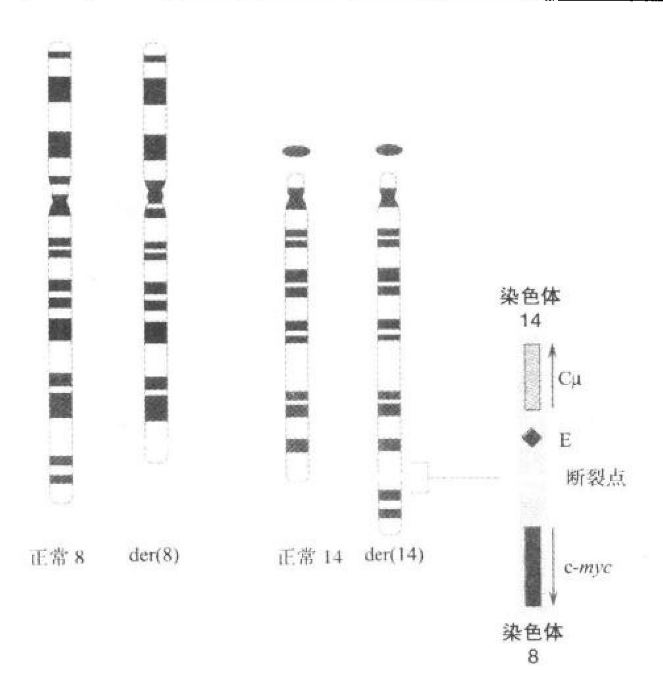


图 11-9　在 Burkitt 淋巴瘤中见到的特异易位。这类肿瘤的大多数表现为易位，断裂点位于 8q24 和 14q32。这些易位的分子水平分析显示，8 号染色体上的断裂点毗邻或位于 *MYC* 原癌基因内；而 14 号染色体上的断裂点在免疫球蛋白的重链基因座内。这种重排的确切性质因肿瘤而异，但本图右侧部分显示了一个典型的例子。在某些情况下，重链基因(*Cu*)的增强子(*E*)可能在激活 *MYC* 基因的过表达中发挥作用。

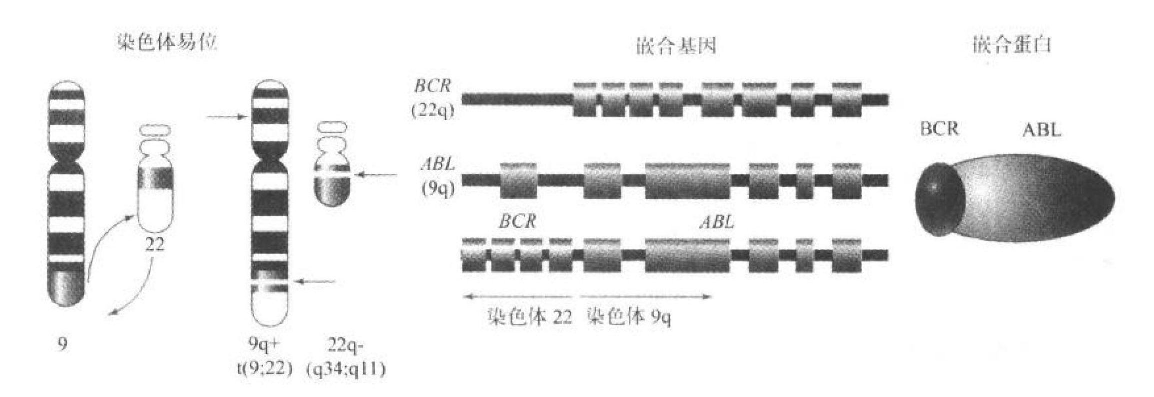


图 11-10　费城染色体易位将 *BCR* 基因的 5′端(在染色体 22 q 上)和 *ABL* 基因的 3′端部分(在染色体 9q 上)相融合。由此产生的嵌合基因将指导嵌合性融合蛋白的表达，该蛋白含有 *BCR* 基因的 N 末端部分，后者与 *ABL* 基因的活性蛋白激酶部分相连(引自 Fearon ER . Oncogenes and tumor suppressor genes. In: Abeloff MD, Armitage JQ, Lichter AS, Niederhuber JE, eds. Clinical oncology. Churchill Livingstone, 1995.)。

(五) 激活的癌基因可以引起家族性癌吗?

以上所述引起细胞原癌基因激活的所有突变是癌细胞中获得的(体细胞的)突变。就是说,当检测患有这些癌症患者的 DNA 组成时,未发现这些突变,表明是在恶变发生过程中获得了突变。然而,正如上述,在真正恶性转化发生之前,似乎很可能需要一次以上这样的突变事件。这就提出一个问题,是否某些家族性癌综合征是由于生殖系遗传激活的癌基因所致。这样的个体如果发生一次额外的体细胞突变,将极易发展为恶性。实际上,通过在适当的靶组织中表达多种癌基因,已使转基因小鼠产生肿瘤。然而,对人类的遗传性癌症而言,这可能是非常罕见的机制。尽管有若干重要的人类遗传性癌症易感综合征,但它们通常涉及到另一类基因的遗传性突变,这类基因称为肿瘤抑制基因或抑癌基因(tumor suppressor genes),我们将在下一节讨论。尽管原癌基因的突变在散发性癌症的体细胞突变中发挥重要作用,但在家族性癌的生殖系中一般没有该基因的突变,这种现象仍令人困惑不解。后面将讨论一个特例(*RET* 原癌基因)。或许由于突变引起主要细胞过程的扰乱并诱发恶变,所以这些突变所产生的表型易受很强的负性选择。

五、肿瘤抑制基因

正如在开始引文中 Boveri 所指出的,有理由假定可能存在一些正常地限制生长的基因,这些基因的失活将会参与恶变的进展,这是一个极为重要的途径。

(一) 视网膜母细胞瘤和 Knudson 假说

20 世纪 70 年代初期,Alfred Knudson 为解释遗传性视网膜母细胞瘤首次提出了肿瘤抑制基因模型。始于 80 年代中期,一系列精心设计的实验证明 Knudson 假说非常精确。

视网膜母细胞瘤是一种视网膜细胞的肿瘤,它发生于儿童出生后和大约 4 岁之间。暂时人们已知大约半数的视网膜母细胞瘤病例是散发的,而且通常只是单侧眼睛发生一个肿瘤。然而,另外 1/2 病例以常染色体显性方式发生,通常累及双眼,发病年龄较小,且常在每只眼睛中发生一个以上的肿瘤。通常外科疗法是摘除受累的一只或双侧眼球。有趣的是,那些家族性视网膜母细胞瘤的患者以后患其他恶性肿瘤的几率增加,尤其是一种骨的恶性肿瘤——骨肉瘤。图 11-11 表示典型的系谱,它鲜明对比了家族性和散发性的视网膜母细胞瘤。

图 11-12 是一个家族性视网膜母细胞瘤患者眼睛的切片图,可见在两个部位有肿瘤形成。然而,整个眼球内衬一层视网膜细胞,并且经检查表明大多数视网膜细胞的功能是正常的。然而,所有这些细胞一定携带视网膜母细胞瘤基因突变,因为这是通过生殖系由受累的双亲之一遗传而来。因此那些发展为肿瘤的细胞一定获得至少一个额外的遗传异常,从而导致它们的克隆增长。由于视网膜的大多数细胞具有风险,另外一次低频率突变("二次打击")仍将足以把少数这些细胞转变为肿瘤发生的行为。

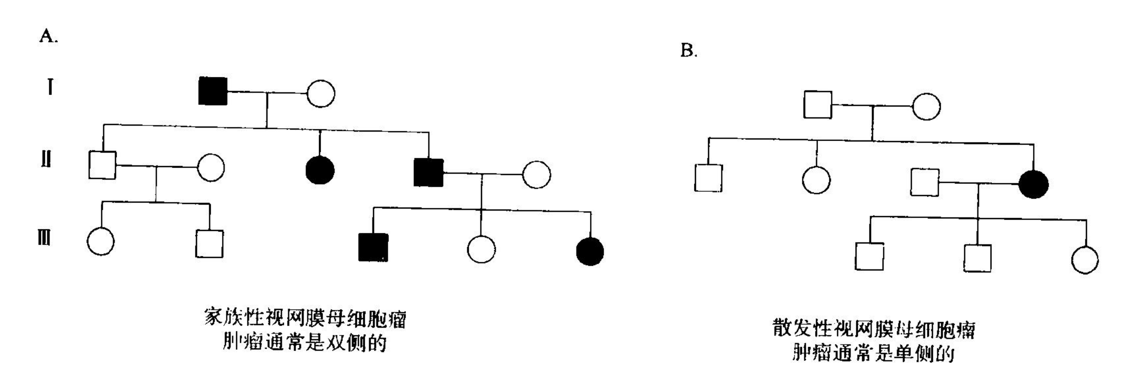


图 11-11 视网膜母细胞瘤的典型系谱。A.家族性显性遗传方式。在这种情况下,肿瘤通常是双侧的,且发生年龄早。B.散发性视网膜母细胞瘤。在这种情况下,某个家庭中只有一个成员受累,且肿瘤发生晚,常常是单侧的。



图 11-12 一个家族性视网膜母细胞瘤个体的眼睛切片图。注意,虽然两个独立的肿瘤很明显,但薄层视网膜的大部分(在制备过程中人为将其与眼的后部分开)是正常的。这表明肿瘤的形成需要在生殖系突变的基础上再发生一次体细胞事件(引自 Willis RA, ed. Pathology of tumors, 3rd ed Butterworth. 1960;894.)。

1971年，Alfred Knudson 提出了如图 11-13 所示的具体假说，将这一观察又向前推进一步。他指出，在家族性视网膜母细胞瘤中的生殖系事件，是一个常染色体肿瘤抑制基因的一个等位基因失活。因为另外一个等位基因仍然保持活性，其蛋白产物的量只减少50%，预期可以忽略这种效应。他又指出在视网膜母细胞瘤中体细胞二次打击是另一个等位基因的丧失。在发生这种事件的细胞中，这种肿瘤抑制蛋白的产生将完全消失，其结果将是生长失控。这一模型得出一个自相矛盾的结论：虽然发生视网膜母细胞瘤的遗传倾向是显性的，但在特异细胞中肿瘤发生的真实机制是隐性的，因为只有视网膜母细胞瘤基因的两个拷贝都失活时才能发生肿瘤。因此，肿瘤抑制基因有时也称作隐性肿瘤基因。并认为这个模型的散发病例是由于在单一视网膜细胞中两个等位基因都发生体细胞的失活，而不存在生殖系突变。

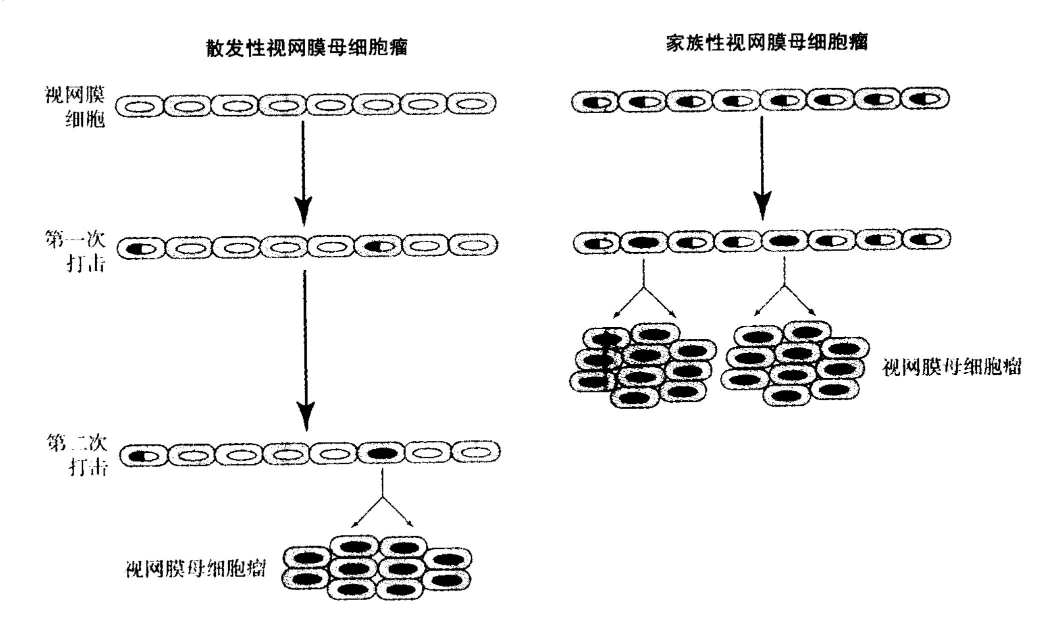


图 11-13 Knudson 关于视网膜母细胞瘤发生的"二次打击"模型。左侧正常的视网膜由一层细胞组成，其 *RB1* 基因的两个拷贝均正常(空心核)。罕见的体细胞突变事件引起少数细胞携带 *RB1* 基因的一个突变("第一次打击")拷贝(半实心核)。如果一个罕见的"二次打击"发生在这些细胞之一，*RB1* 基因的另一个拷贝失活(实心核)，这导致散发性视网膜母细胞瘤。如右侧所描述的家族性视网膜母细胞瘤患者中，所有的视网膜细胞中 *RB1* 基因的一个拷贝在出生时已失活，因此只需要"一次打击"就可产生视网膜母细胞瘤。

(二) 肿瘤抑制基因假说的证据

支持 Knudson 假说的观察之一，是发现极少数患有视网膜母细胞瘤的儿童同时有多种出生缺陷和智力低下，这些是视网膜母细胞瘤的散发病例，但肿瘤却发生在双侧。对这些患者的大多数进行详细的细胞遗传学分析表明，13 号染色体长臂存在中间缺失。图 11-14 展示了所发现的一些缺失。然而，13 号染色体缺失的确切部位却因人而异，13q14 的一个区域在所有的患者中都是常见的。因此这一区域很可能是视网膜母细胞瘤基因

的候选区域。这些细胞遗传缺失的异常患者可能由于染色体的缺失已丢失了视网膜母细胞瘤基因的一个拷贝;染色体的缺失较大也足以累及许多其他基因,这或许可能解释这些患者其他的出生缺陷及智力低下(相邻基因综合征的一个例子,见第八章)。

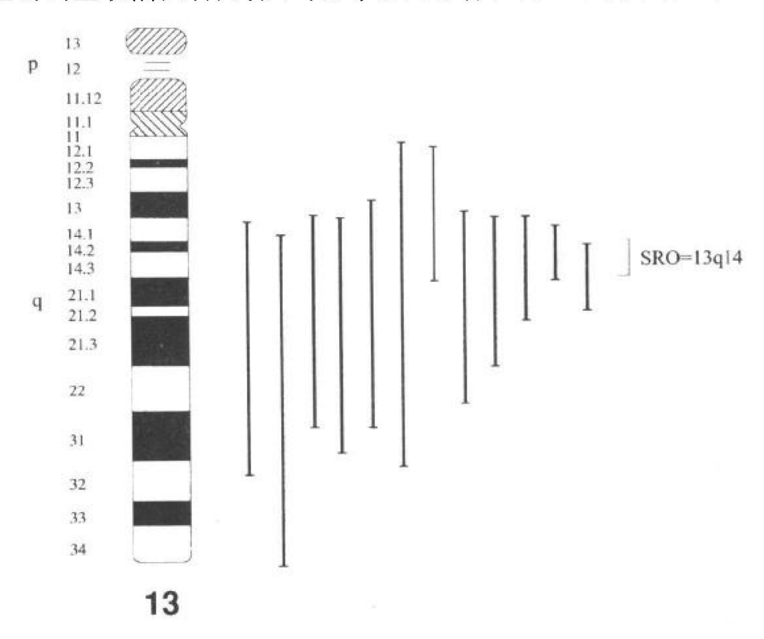


图 11-14　在视网膜母细胞瘤、出生缺陷及智力低下患者中鉴定的 13 号染色体缺失的代表性例子。虽然 13 号染色体缺失的确切位置因患者而异,但所有这些患者都共有 13q14 缺失,因此该区域可能含有视网膜母细胞瘤基因。

如果视网膜母细胞瘤基因位于 13q14,那么人们有望在这种遗传性疾病的患者中发现与这个区域的连锁。对许多视网膜母细胞瘤家系的分析确实发现了这样的连锁,支持这种观点,即在家族性视网膜母细胞瘤中同一个基因(即 *RB1*)发生突变,并且在具有 13q14 缺失的极少数患者中该基因完全缺失。

视网膜母细胞瘤的遗传分析

虽然上述资料支持视网膜母细胞瘤基因定位于 13 号染色体,但有必要对肿瘤组织本身进行分析以证实肿瘤形成机制与该基因的正常等位基因丢失有关。这种丢失的发生可有几种机制, 如图 11-15 所示。最简单的是整条 13 号染色体的丢失。另一种情况是正常的 13 号染色体可能丢失,然后异常的 13 号染色体复制。一种更复杂的机制涉及有丝分裂过程中的重组,即在有丝分裂过程中同源染色体发生交换(值得注意的是,虽然在减数分裂中这是一种基本事件,但直至对这种肿瘤进行研究后才清楚地认识到它在有丝分裂中也存在)。最后,正常染色体上的视网膜母细胞瘤基因可能独自发生不同的突变。这些事件的结果以及它们对与 *RB1* 基因邻近的多态性 DNA 标记的效应如图 11-15 所示。除了独立突变以外的所有这些事件将导致肿瘤组织中一个等位基因及其紧密连锁的 DNA 标记丢失。特别是,如果这些机制之一发挥作用,可以推测在肿瘤中携带正常 *RB1* 基因的染色体上的等位基因将丢失。这即所谓"变为纯合"或**杂合性丢失**(loss of het-

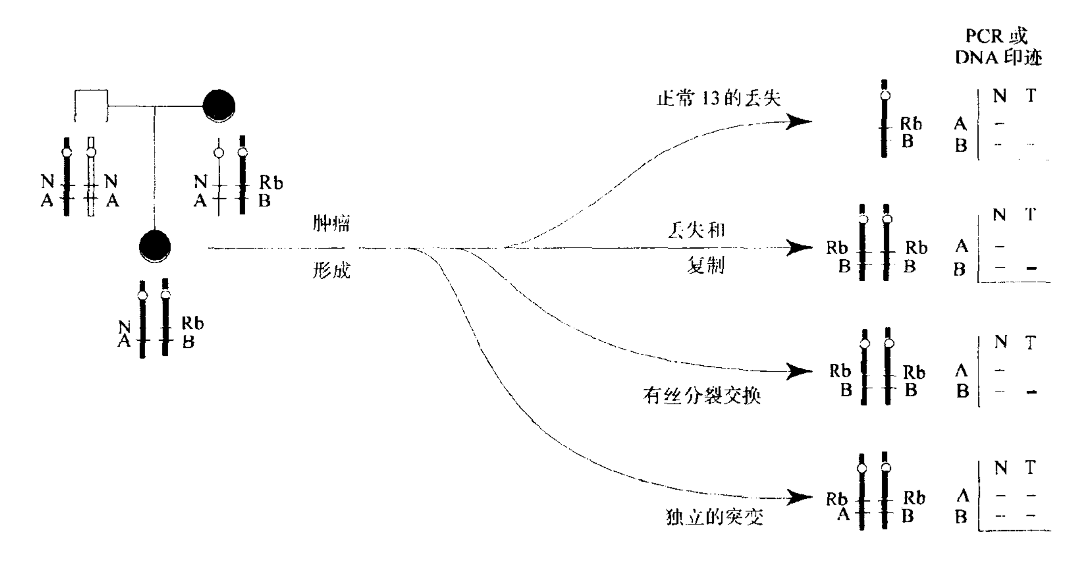


图11-15 家族性视网膜母细胞瘤个体肿瘤形成的可能机制图解。左侧显示的是受累个体的系谱，她从其受累母亲遗传了异常的(*Rb*)基因。画出其受累双亲的四条染色体以表明它们的来源。也分析了此家系中恰在视网膜母细胞瘤基因座下游的一个多态标记，同她的母亲一样，患者在此基因座是AB，而其父亲为AA。因此，B等位基因一定位于携带视网膜母细胞瘤致病基因的染色体上。当这个患者从她父亲遗传的正常等位基因(N)失活时，即发生肿瘤。在右侧列举了这种事件发生可能的四种方式。在每种情况下，都可见由此引起的13号染色体重排。另外用PCR或DNA印迹方法对比分析了正常组织和肿瘤组织的多态标记。注意：在前三种情况下肿瘤组织中正常等位基因(A)均丢失，这称为杂合性丢失(LOH)。

erozygosity，LOH)，这可以通过PCR或DNA印迹方法分析比较视网膜母细胞瘤患者结构性DNA与肿瘤的DNA而检测出来。

由于正常的染色体完全丢失，伴有或不伴有异常染色体的复制，注意在那条染色体上的所有DNA标记将表现为LOH。然而，有丝分裂重组将导致重组位点远端的所有标记的LOH。如果在不同的肿瘤中能检测到几个这样的事件，就更可能对肿瘤基因位于什么位置有一个初步的看法。因为在重组事件中它一定总是要变为纯合性。

仔细分析视网膜母细胞瘤显示了所有这些可能性的发生。通过研究家系，可证实在肿瘤形成中常有正常13号染色体丢失。因此，这些资料与Knudson假说的预测非常吻合，所以有力地支持在染色体13q14上存在一个肿瘤抑制基因。对视网膜母细胞瘤患者中后来发生的其他肿瘤如骨肉瘤的分析通常表明，它与正常13号染色体等位基因丢失具有同样模式，正是由于这些同样的机制之一所致。

(三) 视网膜母细胞瘤基因的克隆

利用对家族进行遗传性分析及对大量肿瘤进行LOH的研究这一优势，于1986年克隆了*RB1*基因，这是首次成功应用定位克隆方法的实例之一(第九章)。正如Knudson假说所预测的那样，利用这种cDNA作为探针对视网膜母细胞瘤进行RNA印迹分析表明，

在某些肿瘤中其信号完全缺失，而在另一些肿瘤中则可见异常大小的信号（图 11-16）。RNA 印迹分析表明一些肿瘤的 mRNA 似乎是正常的，但更详细的测序分析显示其存在点突变。进行细致的分析也没有在视网膜母细胞瘤中发现该基因的正常转录物。

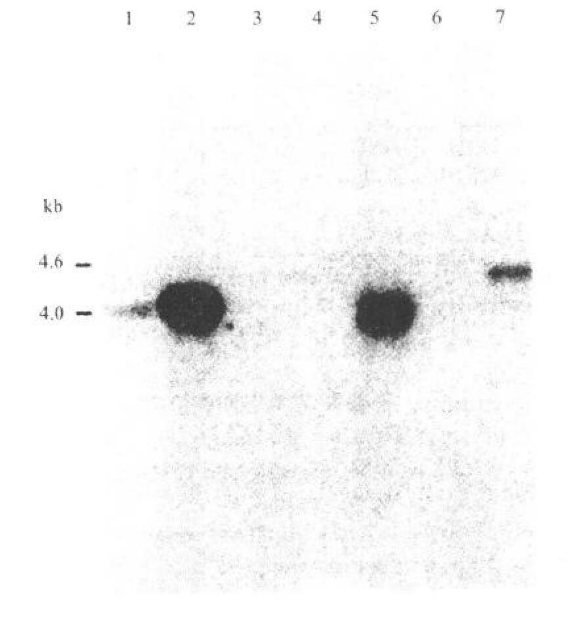


图 11-16 多例视网膜肿瘤 RNA 的 RNA 印迹。RNA 印迹所用探针是代表 *RB1* 基因转录物的 cDNA 克隆。第 1～6 泳道代表来自 6 个不同视网膜母细胞肿瘤的 RNA 标本，而第 7 泳道为来自正常胎儿视网膜的 RNA。注意：该基因正常转录物的长度是 4.6kb，而视网膜母细胞肿瘤的 RNA 或者完全缺乏此转录物（第 3,4 和 6 泳道），或者含有缩短的异常转录物（第 1,2 和 5 泳道）。因此，没有一个视网膜母细胞肿瘤含有正常转录物。这表明在这些肿瘤中，*RB1* 基因座的两个等位基因都是异常的，与 Knudson 模型的预测一致（引自 Lee WH, et al. Human retinoblastoma susceptibility gene: cloning、identificationand sequence。Science 1987;235:1394.）。

自从发现了 *RB1* 基因以来，人们作出了相当大的努力以明确该基因的正常功能。遗传性的 *RB1* 基因突变与视网膜和骨肿瘤的特异相关性显示了 *RB1* 基因在这些组织中的重要功能。因此，观察到 *RB1* 基因在全身许多组织中广泛表达是出人意料的。RB1 蛋白位于核内，可能在细胞整个周期的进程中发挥中心的调节作用。在细胞内 *RB1* 与一些其他调节蛋白相互作用，反过来却又受一系列称为细胞周期素的细胞周期特异蛋白控制。这个过程调节决定一个细胞是进行分裂、增殖还是保持静止。为什么 *RB1* 基因功能异常特别易患视网膜和骨的癌症，仍是不解之谜。正如我们将要看到的那样，与特殊类型肿瘤相关的许多其他基因也在多种细胞共同的关键过程中发挥作用。尽管癌遗传学取得了相当大的进展，但对特异基因缺陷与特殊类型肿瘤间关系的了解还远远不够。

RB1 基因的鉴定是基础细胞生物学重大进展的一个例子，它是对一种罕见的人类疾病研究的结果。这是医学科学史上经常重现的一个经验。

(四)其他罕见的家族性癌综合征

自从发现了视网膜母细胞瘤基因以来,已应用类似的方法鉴定了一些其他罕见的家族性癌综合征相关的基因,如表 11-1 所列出。对于多数但不是所有的这些基因而言,在相关肿瘤中都鉴定了 LOH,这说明 Knudson 最初提出的二次打击模型对许多肿瘤抑制基因都是适用的。正如 *RB1* 基因一样,若干这些基因已证明是一般细胞结构的主要成分。例如,位于3号染色体上的Von Hippel-Lindau基因(*VHL*)与几种高度保守的转录延长因

表 11-1 家族性癌综合征和相关的肿瘤抑制基因

家族性癌综合征	肿瘤抑制基因	功 能	染色体定位	观察的肿瘤类型
家族性视网膜母细胞瘤	*RB1*	细胞周期调节	13q14	视网膜母细胞瘤,成骨肉瘤
Li-Fraumeni 综合征	*P53*	调节细胞周期和凋亡	17p13	脑肿瘤,肉瘤,白血病,乳腺癌
家族性腺瘤样息肉	*APC*	信号从细胞表面黏附分子向核传导	5q21	结肠癌
Von Hippel-Lindau 综合征	*VHL*	调节转录延长	3p25	肾癌,成血管细胞瘤,嗜铬细胞瘤
Wilms 瘤	*WT1*	转录调节	11p13	儿童肾癌
家族性黑色素瘤	*CDKN2A*	细胞周期调节	9p21	黑色素瘤,胰腺癌,其他
遗传性非息肉性结肠癌	*MSH2*	DNA 错配修复	2p16	结肠癌
遗传性非息肉性结肠癌	*MLH1*	DNA 错配修复	3p21	结肠癌
家族性乳腺癌	*BRCA1*	?	17q21	乳腺和卵巢癌
家族性乳腺癌	*BRCA2*	?	13q12	乳腺和卵巢癌
Gorlin 综合征	*PTCH*	? 早期发育和细胞分化	9q22	基底细胞皮肤癌
多发性内分泌肿瘤Ⅰ型	*MEN1*	?	11q13	甲状旁腺和垂体腺瘤,胰岛细胞肿瘤,类癌瘤
多发性神经纤维瘤Ⅰ型	*NF1*	催化 RAS 失活	17q11.2	神经纤维瘤 ,肉瘤,胶质瘤

子相互作用,可能在调节正常 RNA 转录过程中发挥重要作用。遗传了 *VHL* 基因生殖系突变的患者发生多发性肾癌,以及视网膜、小脑及脊髓的异常血管肿瘤(成血管细胞瘤,hemangioblastoma)。在大多数散发性肾癌中都可鉴定 *VHL* 基因的体细胞突变。已证实在家族性黑色素瘤中,存在与细胞周期调节有关的特异基因如 *CDKN2A* 的突变。另外一个例子是,最近鉴定的一个基底细胞痣综合征(或 Gorlin 综合征——它是一种与多种基底细

胞皮肤癌和许多出生缺陷有关的少见遗传病)的基因,它与一种称为 *patched* 的基因密切相关,*patched* 基因调节黑蝇属果蝇早期胚胎发育和分化。

显然,从我们至今所讨论的例证,家族性癌综合征的主要基因通常是肿瘤抑制型,在细胞水平以隐性方式发挥作用。然而,综合征的遗传方式为常染色体显性。一个原癌基因遗传性激活突变的罕见例子见于 *RET* 基因。导致多发性内分泌肿瘤 2 型的形成(Multiple endocrine neoplasia type2 ,MEN2)。*RET* 原癌基因最初是经 DNA 转染实验鉴定为癌基因的,它是跨膜酪氨酸激酶生长因子受体大家族成员之一。现已发现约 30% 滤泡性甲状腺癌中 *RET* 基因存在重排,并且通过靶向诱导甲状腺 *RET* 突变建立了这种疾病的小鼠模型。在某些乳腺癌中还发现该基因的扩增。MEN2 患者通常带有 *RET* 基因遗传性生殖系突变,并引起其受体成分的激活。MEN2 患者发生的特征性甲状腺癌称做髓性癌(medullary carcinoma),也发生肾上腺和甲状旁腺肿瘤。虽然在 MEN 相关的肿瘤中许多染色体可见 LOH,但这些位点不包括 *RET* 基因座本身。奇怪的是,*RET* 基因功能丧失突变产生完全不同的疾病,称为 Hirchsprung 病,表现为结肠和直肠缺少正常自主神经丛,导致结肠明显扩张。这是同一基因的不同突变可产生截然不同表型的一个明显的实例。

P53 基因提供了另外一个很有说服力的实例。即一个基因既可以是肿瘤抑制基因,又可以是一个显性作用的癌基因。NIH3T3 细胞转染实验表明,*P53* 可发挥癌基因功能使细胞永生化,也可与突变的 *RAS* 基因协同转化细胞。后来又证实在许多类型肿瘤中,*P53* 基因突变是最常见的体细胞突变之一。P53 蛋白是一转录因子,它在调节细胞的细胞周期进程以及细胞进入凋亡(程序性细胞死亡,programmed cell death)的途径中发挥关键性作用。

现在看来,正常 *P53* 基因主要作为一种肿瘤抑制基因发挥作用。*P53* 功能的丧失解除了对细胞生长的正常抑制,使突变细胞获得生长优势。*P53* 功能的丧失也使细胞逃脱凋亡。*P53* 基因生殖系突变导致常染色体显性癌症易感性疾病,Li-Fraumeni 综合征。就可以观察到的肿瘤多形性分类而言,这种疾病在遗传性癌综合征中非常显著。患者常特异地发生各种软组织肉瘤以及乳腺癌、脑肿瘤、白血病和许多其他类型癌症。这些肿瘤也较早发生,50% 患者在 30 岁时就发生第一个癌症。

六、DNA 修复基因缺陷

至今我们已讨论的所有遗传性癌易感综合征,以及我们将在本章后部分将讨论最近鉴定的与一般癌症有关的基因,是以常染色体显性方式遗传的。目前仅有少数已知的隐性癌症易感综合征。一些这类综合征的遗传缺陷涉及的基因与 DNA 修复或复制有关。例如,某些不同类型的着色性干皮病(xeroderma pigmentosium)是由于修复紫外线对 DNA 损伤的基因突变所致。这些患者发生多发性皮肤癌,尤其是暴露于阳光(紫外线)的部分,也就不足为奇了。Fanconi 贫血和 Bloom 综合征的遗传缺陷引起基因组不稳定和频繁的染色体断裂。在这些综合征中除了癌症发生频率高以外,这些患者也常有某些特征性的骨骼和其他发育异常。

最近,通过对遗传性结肠癌的研究,已鉴定了与 DNA 错配修复有关的一类新基因。以下我们将集中讨论这组“增变基因”(mutator genes),以明确癌症发生途径中的一种新机制。

七、常见癌症的遗传易感性

直到最近，认为遗传类型的癌只限于一些罕见的综合征。尽管相同的基因可通过偶发的体细胞突变事件参与散发性癌的进展，但一般认为生殖系突变与常规医学实践的直接相关性是有限的。由于过去几年中，在很常见的两种癌症——乳腺癌和结肠癌的大量患者中鉴定了特异性生殖系突变，这种情况已发生显著的变化。最近也定位了前列腺癌的一个易感基因。尽管其他一些常见癌症如肺癌（死亡的首要原因）的遗传易感性还未得到证实，但这种情况在未来几年内可能会发生改变。

对罕见的家族性癌综合征进行的某些一般观察为认识常见癌症的遗传类型提供了有益线索。由生殖系突变所致的癌症发生年龄一般较散发性癌症年轻，并且在同一个体常产生多发性肿瘤。对与其他不常见类型癌症的相关性也可提供有益线索。然而，通过在一般群体中不相关散发癌症的大背景鉴定常见癌症易感基因相当困难。例如，乳腺癌是妇女中最常见的癌症，在她们的一生中受累者为11%。尽管现在看来5%～10%的乳腺癌可能与几种特异易感基因之一的生殖系突变遗传有关，但大多数病例（约95%）是散发的，这可能通过偶发体细胞突变、环境因素的参与和多个微效易感基因间的相互作用所致（参见第四章多因素疾病的讨论）。偶尔，散发性乳腺癌先证者姐妹的11%将发生散发性乳腺癌。同样，散发性乳腺癌患者的姐妹和母亲均独立发生散发性乳腺癌的可能性大约为1%。由于这种原因，很难区分乳腺癌是由一个癌症基因遗传引起的，还是在一个家系中由散发性乳腺癌病例聚集所致。这一问题也为癌家族的临床管理提出一个重要的挑战。这种情况将与罕见肿瘤如MEN2相关的家族性癌综合征形成对比。同一家族的一些个体中发生不常见的甲状腺髓性癌和嗜铬细胞癌，绝不可能是偶然的。同样，在与视网膜成血管细胞瘤（一种在任何其他群体中罕见的肿瘤）有关的von Hippel-Lindau综合征中发生的肾细胞癌，使散发性肾细胞癌很容易与此综合征相关的那些疾病区分开来。

八、家族性乳腺癌

家族史长期被认为是乳腺癌发生最重要的危险因子。然而由于我们刚刚概述的原因，很难将一种明确的常染色体显性遗传乳腺癌与该病很常见的散发型区分开来。结果，直到20世纪80年代末期，才发现一种明确的常染色体显性乳腺癌易感基因。通过定位克隆方法现已鉴定了两个这样的乳腺癌易感基因，即*BRCA1*和*BRCA2*。尽管至少可能有另一个主要的*BRCA*基因有待鉴定，但是这两个基因一起可能是所有乳腺癌中多达5%～10%的病例的原因。虽然在受累家系中LOH的证据表明*BRCA1*和*BRCA2*基因是肿瘤抑制基因，但对它们在细胞内的功能仍知之甚少。遗传*BRCA1*或*BRCA2*基因生殖系突变的妇女在其一生中患乳腺癌风险为60%～90%，患卵巢癌的风险为20%～60%。在*BRCA*家族中患其他癌症的风险也可能出现轻度增加，包括*BRCA1*引起前列腺癌、*BRCA2*引起男性乳腺癌。*BRCA*基因的鉴别已经提出一个难题，即DNA检测和突变筛查在患者临床管理中的作用。此问题将在本章后面加以讨论。

九、家族性结肠癌

家族性结肠息肉病是一种罕见的常染色体显性遗传的结肠癌(发生率约 1:10 000),通常很容易与散发性结肠癌相区别。虽然家族性息肉病患者在出生时结肠是正常的,但在生后的前 20 年中其结肠上或偶尔在肠道的其他部位上可出现数百个小息肉。尽管这些息肉没有症状,但它们的主要意义在于有发展为结肠癌的风险,这些疾病的患者到 50 岁时,100%将发生结肠癌。在成年早期进行全结肠切除可完全防止这种结果。图 11-17 为此病患者的 X 线照片和病理标本,可见无数个息肉。因为这是一种常染色体显性疾病,在恶性转化出现以前没有其他警告信号反映该病的存在,所以在一个患者作出诊断后,医生应检查家族的其他成员。因为外科治疗如此成功,所以有风险的个体应在 20 岁前通过 DNA 检测,如需要可做结肠镜以进行评价。

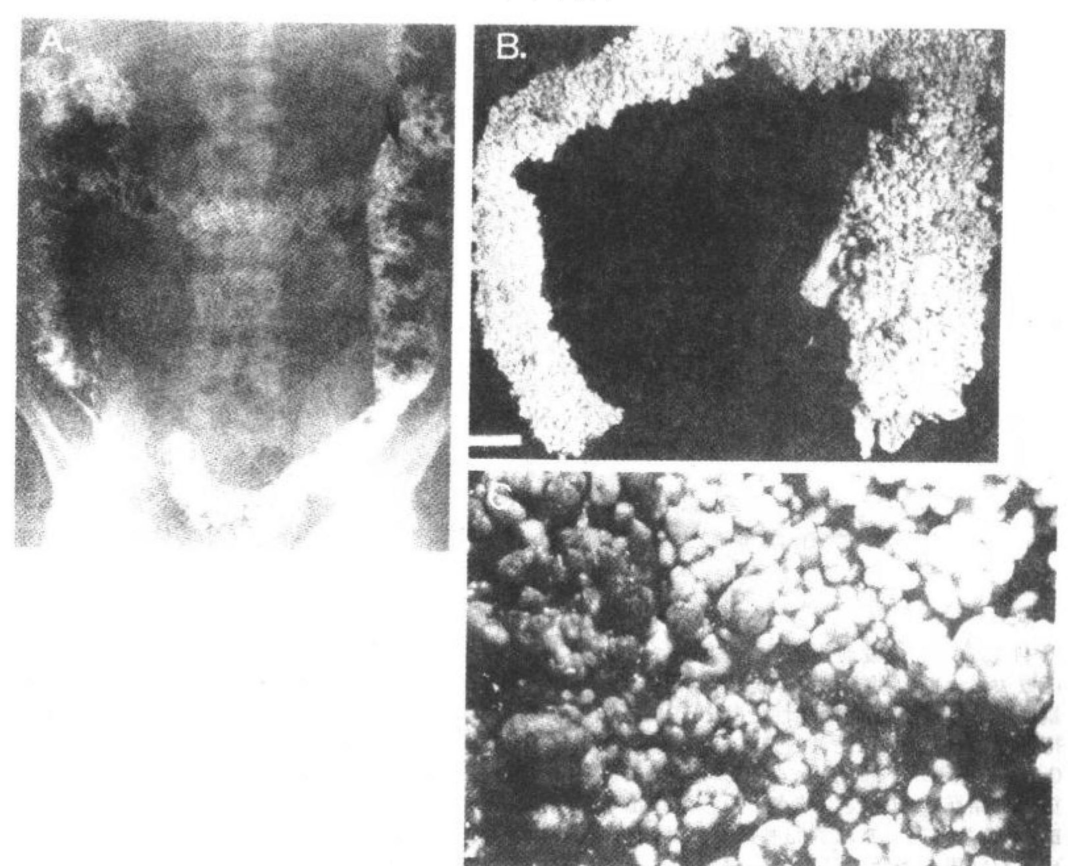


图 11-17 家族性结肠息肉病。A.受累患者的钡灌肠图。结肠充满了放射密度的钡,大肠内的无数息肉表现为肠腔内充盈缺损。B.从受累患者切下的结肠,表明多个息肉呈毯状分布于整个大肠。C.一个切下标本的直观图。由于每个息肉都有发展为结肠癌的风险,这就是为什么全结肠切除是惟一成功治疗方法显见的原因(引自 Sleisenger MH, Fordtran JS. Gastointestinal disease : pathophysiology , diagnosis ,management.4 th ed. Philadelphia: WB Saunders Co.1989:1502)。

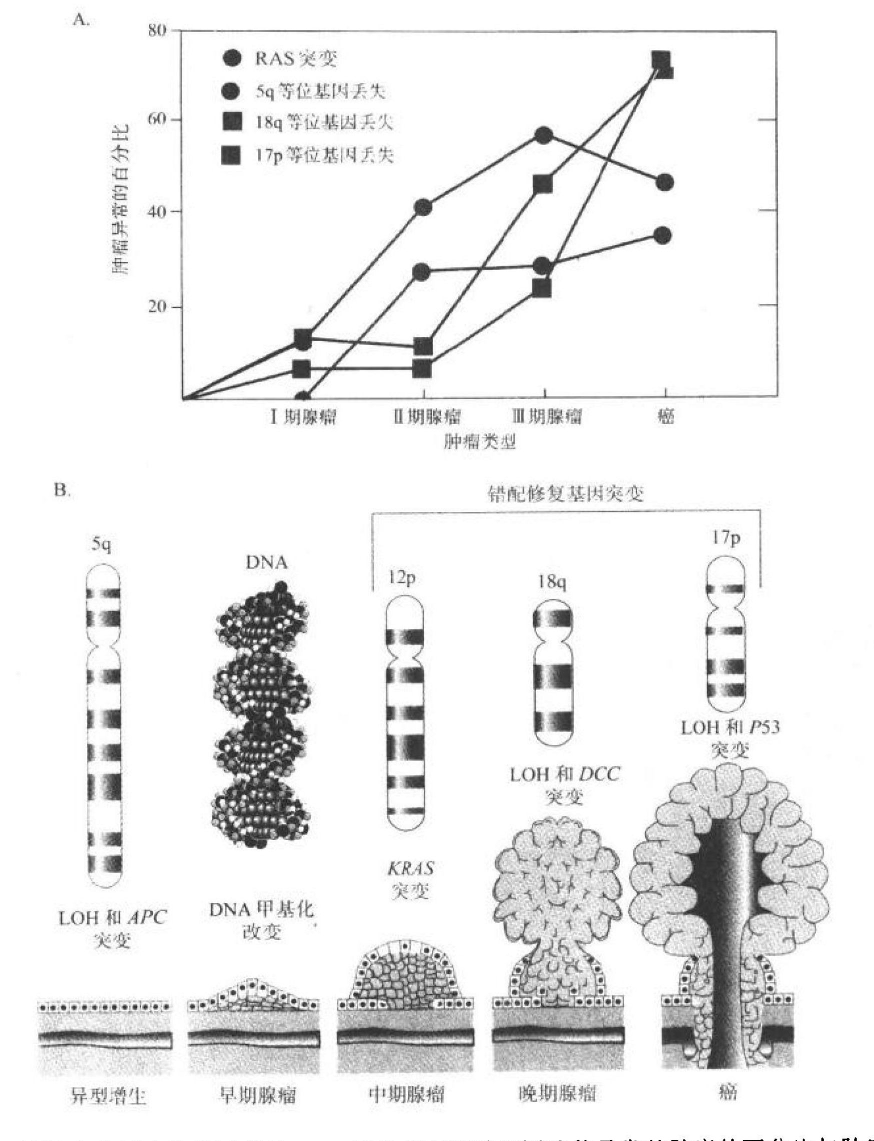


图 11-18 结肠癌发展中的分子事件。A. 显示具有四种不同遗传异常的肿瘤的百分比与肿瘤类型呈函数关系。I型腺瘤，最良性的，但也有最低的遗传异常发生率。随着恶性度的增加，突变事件的数量也随之增多。这包括显性癌基因(*RAS*)的激活，以及在5q(*APC*)，18q(*DCC*)和17p(*P53*)上的肿瘤抑制基因的失活，这是通过这些基因座的LOH推断的(引自 Vogelstein B, FearonER, Hamilton SR, et al. Genetic alterations during co;orectal tumor development. N Engl J Med 1988;319:525.)。这些观察为B中所描述的结肠直肠癌发生的多阶段模型作出了贡献。

B. *APC*基因突变被认为是肿瘤发生始动过程的早期事件。遗传*APC*基因生殖系突变的家族性腺瘤样息肉病患者经历了这个始动过程的加速，并发生多灶性表皮发育异常，其中许多进展为早期腺瘤或息肉。这种良性肿瘤逐步进展为成熟的癌症是若干体细胞突变事件的结果，包括*RAS*原癌基因的激活，肿瘤抑制基因如*DCC*和*P53*的进行性丢失。错配修复基因突变加速了体细胞突变的过程。无疑在此过程中还有其他一些遗传的改变仍有待证实，一般说来，在此途径中更恶性的肿瘤已获得大量的突变事件。组织学上通常可以看到一种癌是从一个较低度恶性肿瘤的表皮区域发生，并且可以见到恶性度更高的区域已获得另外的遗传事件(引自 Weinberg RA, Hanahan D。The molecular pathogenesis of cancer . In: Bishop JM. Weinberg RA, eds. Scientific American molecular oncology. New York: Scientific American. Inc. *1996*:187.)。

虽然家族性腺瘤样息肉病基因，即 *APC* 基因，于 1991 年首次被克隆，但直至最近才开始出现关于其功能的线索。*APC* 基因可能通过与细胞黏附有关的蛋白家族的相互作用而发挥功能，调节随后的信号向核内传导。*APC* 基因符合经典的肿瘤抑制基因的标准，由于第二次打击使家族性息肉病患者保留的正常等位基因的表达完全丧失。结肠黏膜细胞 *APC* 基因表达的丧失产生克隆性良性肿瘤增殖，进而形成息肉。尽管这个单一事件本身并不能产生癌症，但它却可能是使细胞对其他遗传事件作用显著易感的开端，这些遗传事件的累积最终产生真正的结肠癌。这种机制也可能在散发性结肠癌中发挥作用，至少 80% 散发性肿瘤表现为 *APC* 基因座的体细胞突变。

这些和其他的结肠癌研究所得出的结果，或许为从一个正常细胞进展为完全的癌所需要的多遗传阶段提供了最清楚的例证。对大量结肠癌进行的 LOH 研究验明了在染色体 5q 、17p 和 18q(见图 11-18A)上存在一个连续遗传丧失的模型。最早的事件是涉及染色体 5q 上的基因，即后来证明的 *APC* 基因，在 17p 上的是我们已经讨论过的 *P53* 基因，18q 上的基因即 *DCC*“结肠癌中缺失的”(deleted in colon cancer)功能仍然未知，而且尚未鉴定该基因的生殖系突变。结肠癌细胞 DNA 发生其他类型的变化或改变，如 DNA 甲基化，它也可能与恶性表型有关。最近，对与息肉病无关的家族性结肠癌进行的研究已鉴定一组新的增变基因，它们可能也参与散发性结肠癌和其他类型癌症的发展。

十、遗传性非息肉性结肠癌

虽然家族性息肉病是不到 1% 的结肠癌的原因，但经常可见非息肉结肠癌的家族性聚集。这些家族常具有与家族性癌综合征相关的特征，如发病年龄轻，且与其他特异的肿瘤如卵巢癌和子宫内膜癌有关。这种综合征称为遗传性非息肉性结肠癌(Hereditary nonpolyposis colon cancer, HNPCC)或 Lynch 综合征。应用现在熟知的定位克隆方法，于 1993 年首次鉴定了与多数 HNPCC 家族发病有关的基因，称为 *MSH2*。该基因和细菌及酵母体内参与 DNA 序列错配修复的一个基因家族密切相关。后来在 HNPCC 家族中又鉴定了其他三个基因，也可能参与同样的 DNA 修复过程。与此功能一致的是，HNPCC 患者的结肠癌细胞特别容易自发突变。这种改变或许增加其他基因额外体细胞突变的发生，因此参与由正常细胞进展为癌细胞的多阶段过程(如图 11-18B)。这些额外的体细胞突变的特异靶基因正是目前加深研究的一个课题。

尽管 DNA 错配修复基因似乎是大多数 HNPCC 家族的病因，但仍有待鉴定另外一些基因。在一般群体中，HNPCC 基因生殖系突变的发生可高达 1:200，可能是多达所有结肠癌的 10% ~ 15% 的病因。不管涉及什么基因，HNPCC 家族间相似的表型为基因座异质性提供了极好的例证(见第三章)。

十一、癌的一般机制

从本章介绍的资料可得出一些原理。Boveri 在开篇引言中提出遗传上以这样一种方式改变一个细胞使其生长增加的两种主要机制。事实上，如果他用“基因”一词来替代

"染色体",则那一段落准确地描述了癌基因和肿瘤抑制基因。

我们已看到在正常细胞转化为恶性细胞的进程中,原癌基因激活是一常见的体细胞事件。肿瘤抑制基因通常需要二次"打击"才失活。虽然两个等位基因都可以通过体细胞事件失活,但遗传了一个已经失活的等位基因的个体表现出显著的癌症易患性。第三类增变基因直接干预 DNA 修复过程,导致更普遍的遗传不稳定性。

一个正常细胞转化为恶性细胞需要一次以上的突变事件的事实,是这些讨论中一个反复出现的主题。图 11-18B 描述了一个正常上皮细胞逐渐进展为成熟的结肠癌需要多个不同的突变事件。

这一多阶段模型作为癌症起源的一般范例如图 11-19 所示。左侧开始是一个正常细胞,在适当条件下它有抑制其自身生长所必需的机制。单个突变可以稍微增加那个细胞的增殖能力,但是尚不足以完全逃脱正常的调控机制。为什么一定如此,事实上存在着一些不得不相信的理由。假定体内约有 10^{14} 细胞,估计在复制期间每个基因的突变率为 10^{-6},如果一单个突变足以致癌,那么癌症将是个日常事件。单个突变的细胞可具有完全正常的表型或轻微的增长优势,使其得以存活、繁殖,并最终获得另一个突变。经历这一途径的每一阶段后,细胞获得更大的自主性,因此比其相邻的细胞能更好地繁殖。最终达到一点,即此时细胞可以不受正常时阻止其增殖的信号的影响而生长,然后它可继续破坏邻近组织并扩散到身体的其他部位。

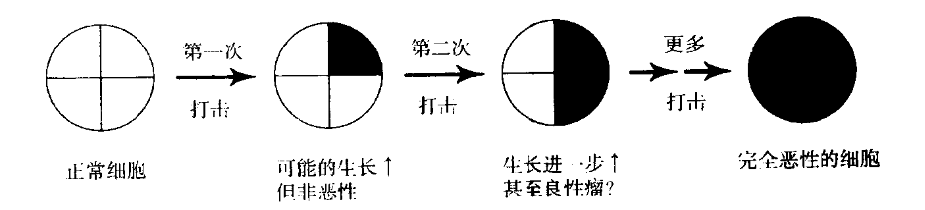


图 11-19　癌症多阶段发生的图解。将一个细胞转变成有肿瘤表型的细胞,可能至少必须有两个遗传事件,而且对于某些组织可能需要两个以上的事件。生殖系的癌症易患性可认为是将此图向右推进一步。

代表生殖系癌症易患性的家族性癌综合征被认为是将此图向右推进一步。具有这种生殖系突变的个体仍然是由行为上基本正常的细胞组成。然而,它们比正常情况下向恶变更进一步,并具有更高的风险。

十二、癌症易感性的遗传学诊断

我们在本章已简要概述的癌遗传学的显著进展,几乎全部是在过去的 10~15 年中取得的。虽然实验室研究的重大发现使我们对癌症问题的基本理解发生了革命性变化,并开辟了许多令人兴奋的研究新领域,但对患者治疗的直接影响才显露端倪。人们普遍希望这种科学上的进步能很快地转变为新的治疗方法,包括本章所讨论的许多癌基因的基因治疗。然而,现时这样的新疗法仍是梦想多于现实,承认这一点很重要。目前癌症治疗仍然依赖于外科、化学治疗以及放射治疗的基本原理,这些方法多是在 20 世纪 60 年代

和 70 年代发展起来的。简而言之,癌症遗传学的主要作用基本在于识别癌症风险增加的个体。

许多家族性癌综合征和癌症易感基因的鉴别,展现了预测性 DNA 检测的可能性,这种检测与第五章叙述的应用于其他遗传病研究的一般方法相同。大规模群体筛查 BRCA 和 HNPCC 突变的前景,已在学术界和商业界就新检测技术的开发引起了强烈的兴趣。然而,正如表 11-2 所概括的那样,对癌症易感基因进行广泛 DNA 检测还存在一些障碍。在一个特定家族中,特异的突变通常是独特的,需要对整个基因进行详细的序列分析来证实,这种方法既费时又昂贵,不适用于大规模检测。

即使在一个特定的家族中发现了一个新的突变,有时很难确定是否此突变就是该癌症综合征的病因。对于只引起癌症基因编码序列中单一氨基酸替换的微小突变尤其是个难题。就 *P53* 基因而言,已开发了其功能筛查实验,这取决于野生型 *P53* 基因用基因工程技术导入酵母补偿某种分子缺陷的能力。在这项检测中突变型 *P53* 序列不能发挥作用,这提供了一个强有力的证据,即相应的突变是患者癌症表型的原因。

就 *APC* 基因而言,已鉴定的大约 80%的突变不是无义突变,就是插入或缺失,阻碍全长 APC 蛋白的翻译。基于这种观察,已开发了一种筛查试验(叫做蛋白截短试验,protein trunction test),它可用 PCR 扩增患者血液标本的 mRNA,然后在体外翻译成蛋白质。在这项测定中,一种异常的、截短的蛋白产物的产生为致病性 *APC* 基因突变提供了敏感的筛查方法。

最近通过用第五章所概述的许多 DNA 序列筛查方法,分析了这些基因以及许多其他的癌症易感基因如临床上重要的 *HNPCC* 和 *BRCA* 基因。由于测序技术的不断改进,或许随着自动 DNA"芯片"(chips)的开发,通过杂交快速筛查整个基因,对大量患者进行突变分析的效率会逐渐提高而且费用也能够承受。

癌症综合征症状前 DNA 检测的意义

近期研究已鉴定在特殊限制群体中高频率发生的癌症易感有限数量的突变。例如,两种特异的 *BRCA1* 和一种 *BRCA2* 基因的突变占北欧犹太妇女中 *BRCA* 基因相关乳腺癌的 90%左右。这种明显相关性显然是由于该群体的建立者效应所致。大约 1%北欧犹太妇女携带 185delAG(在第 185 位 A 和 G 核苷酸缺失) *BRCA1* 的突变,同时约 1.4%携带 6174delT *BRCA2* 基因突变。与第三类突变(5382insC)结合起来,估计该人群中 *BRCA* 基因突变的发生率大约是 1:40。相比之下,在一般群体中已鉴定 *BRCA1* 100 余种不同的突变,推测这一群体中 *BRCA1* 基因突变的总人口频率接近 1:500。假定在北欧犹太人中存在有限的 *BRCA* 基因突变,故在此人群中筛查乳腺癌易感性在技术上简单且直接。然而,应用这一简单的筛查实验常带来一些艰难和复杂的伦理问题(表 11-2)。

许多患者关心的主要问题是 DNA 检测结果对他们获得或维持健康保险能力的影响,以及这种信息对未来就业的影响。这些问题引起公众和政界的极大关注,并且将很快促使制定适当的法律以保护患者。

表 11-2 家族性癌综合征普遍 DNA 检测的障碍

技术的
.许多肿瘤抑制基因的大小及复杂性
.群体中存在多个不同的突变
.区别中性错义替换与真正致癌的突变
.高昂的检测费用
医学的
.为获得完全知情同意,需要完善的患者教育
.对大规模检测结果进行遗传咨询的复杂性
.对突变携带者进行最适医疗管理的不确定性
伦理的和社会的
.DNA 检测对健康保险可行性及费用的潜在影响
.失业的可能性
.不能确保检测结果的保密性
.检测对家庭其他成员的影响

DNA 检测结果对不同癌症易感综合征患者的医疗管理的影响明显不同。例如,对于家族性结肠癌患者而言,一个阴性检测结果可使患者免于每年一次费用高且痛苦的结肠镜检查。同样,对 von Hippel-Lindun 综合征而言,通常建议高风险患者每年接受一次 CT 照相和磁共振成像扫描,DNA 检测可使 50% 危险患者免此麻烦。遗憾的是,DNA 检测对医疗管理的影响常不是这样直接。尽管 *P53* 基因生殖系突变(Li-Franmeni 综合征)的患者的癌症风险很高, 多种器官可能受这些癌症影响,使人们难以推荐任何可能有益的标准监视计划。因此,对此疾病进行检测涉及伦理的问题,开始与亨廷顿病症状前检测中的问题很相像,亨廷顿病是另一种尚无特异疗法可行的疾病(见第十四章)。

由于围绕家族性乳腺癌患者的临床管理原则存在相当的不确定性,所以常规筛查 *BRCA* 基因突变目前还存在争议。乳腺 X 线照相是乳腺癌有用的筛查方法,对那些高危人群可在较小的年龄即开始进行,但并不能普遍成功地防止因乳腺癌所致的死亡。尽管预防性乳腺切除术和卵巢切除术降低这些患者乳腺癌和卵巢癌的患病风险,但这些举措也不能完全预防这些癌症。因此,还没有明确的医学管理方面的建议可推荐给那些 *BRCA* 基因突变检测阳性的妇女。在已知 *BRCA* 基因突变的家族中,实验阴性的妇女患散发性乳腺癌的风险仍占总人口的 11%。目前围绕癌症易感性进行的 DNA 检测尚存在相当大的不确定性,医生培训患者并得到患者完全知情同意这一任务具有极大的重要性。

推 荐 读 物

Fearon ER, Cho KR. The molecular biology of cancer . In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics, 3rd ed. New York: Chuchill Livingstone, 1996; 405 ~ 438. *Excellent overview of oncogenes and tumor suppressor genes*.

Bishop JM, Weinberg RA, eds. Scientific American molecular oncology. New York: Scientific American, Inc, 1996.

Nicely illustrated , multi-author review of many aspects of cancer molecular biology and genetics .

Beaudet AL, Scriver CR, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease. CD version. New York: McGraw-Hill, 1997. *This is an outstanding text that covers many aspects of medical genetics and is cited in several other chapters. The recently released CD version contains a new section, edited by KW kinzler and B Vogelstein , with authoritative review of many topics in cancer genetics .*

Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996;87:159 ~ 170. *Excellent review of the current understanding of colon cancer biology by the leaders in the field .*

Tonin P, Weber B, Offit K, et al. Frequency of recurrent *BRCA1 and BRCA2* mutions in Ashkenazi Jewish breast cancer family . Nat Med 1996;2:1179 ~ 1183. *Review of the complex issues of DNA screening for the common breast cancer gene mutations in the Ashkenazi Jewish population .*

Burke W, Petersen G, Lynch P, et al. Recommendations for follow-up care of individuals with an inherited predisposition to cancer Ⅰ. Hereditary nonpolyposis colon cancer. JAMA 1997;277: 915 ~ 919. Burke W, Daly M, Garber J, et al. Recommendations for follow-up care of individuals with an inherited predisposition to cancer Ⅱ. *BRCA1* and *BRCA2*. JAMA 1997;277:997 ~ 1003. *Recommendations for management of patients with hereditary nonpolyposis colon cancer and familial breast cancer from a consensus panel of experts .*

第十二章　临床遗传学

从前，遗传学家像一个“推测者”，为将要发生的任一特定事件预测。现在遗传学家参与“行动”，即诊断和治疗，他已经从一个“推测者”变成“修理师”。

——Roy D. Schmickel（引自 Genetic Counseling As A Form Of Medical Counseling: Univ Michigan Med Ctr J 1974;40:38～43）

随着人们对遗传病的突变机制和分子病理生理学的突破性认识以及人类疾病基因定位的快速进展，遗传学家确实已成为这一行动的中心部分。创建于1979年的美国医学遗传学委员会（The American Board of Medical Genetics, ABMG）于1981年开始给医学遗传学家和遗传咨询家颁发执照。作为成熟的医学专业医学遗传学的重要作用在1991年得到认可，同年美国医学专家委员会（The American Board of Medical Specialties, ABMS）承认ABMG为其成员，这是12年来ABMS承认的第一个新的评审委员会。美国医学遗传学学院创建于1991年，代表医学遗传学的专业团体，并在1996年被美国医学会全国代表大会（The American Medical Association's House of Delegates）接纳为会员。独立的美国遗传咨询委员会（American Board of Genetic Counselors）于1993年创立，其宗旨为评审遗传咨询家。

遗传病患者的临床管理包括很多在其他医学专业中应用的相同的诊断和治疗技术。然而，医学遗传学主要侧重于预防和（或）防止疾病的发生。该方法的核心是症状前诊断、提供遗传咨询、应用遗传筛查项目和产前诊断。其每个题目都将在本章讲述。

一、如何获取家族史及获取家族史的原因

家族史在临床遗传学中有着重要作用。适当地获取和解释家族史是关注遗传病患者的医生可以利用的最有意义和最简便的手段之一。

在考虑如何获取家族史之前，值得考虑的是为什么要获取家族史。首先，家族史有助于获得正确的诊断。第三章中讨论（图3-3）的W.H.在25岁时出现典型的心绞痛胸部疼痛症状，但因其年龄而排除了冠状动脉疾病的诊断。30岁时，W.H.第一次出现心脏病发作，因考虑到他太年轻而不应有心脏病，诊断再次被错过。如果他的医生获得他父亲和叔叔有早发冠状动脉疾病的家族史，他应该被正确地诊断为患家族性高胆固醇血症，并早期进行适当的治疗。

另一个引证的病例是一名因胸部肿块就诊的26岁妇女。她的医生建议她6个月后复诊；这名医生因她年轻而没有认真地考虑乳腺癌发生的可能性。四个月后，她因严重背痛而再次就诊，证明是乳腺癌发生骨转移的结果。随后获得的家族史显示她的一个姐姐26岁时死于乳腺癌转移，另一个姐姐28岁时行乳腺癌双侧扩大切除术。获得了这个

简单的家族史可能会提醒医生认真考虑乳腺癌的诊断并采取适当的检查。

获取家族史的另一个原因是帮助确定正确的预后。R.B 的病例说明了家族史在作出正确诊断、提供适当的预后、提示疾病未来发展过程等方面的意义。由于有五年疑是 Huntington 病的不自主运动(舞蹈症)病史,在 52 岁时她求助于医学遗传学诊所。其运动异常的症状确实像 Huntington 病,但她没有任何痴呆或智力减退的症状。在图 12-1 描述的家族史中,显示她母亲的几个亲属都有相似的疾病,说明她的病是以常染色体显性方式遗传的。然而,引人注目的是没有一个受累个体有任何痴呆症状或出现短寿现象,后两种现象与 Huntington 病的诊断不符,说明该家族中可能存在一种良性舞蹈病,该病在临床严重性和预后方面与 Huntington 病大不相同。随着 CAG 三联体重复扩张在 Huntington 病中的发现,现在该病的诊断可通过 DNA 检测而确认或排除。家族史也提示该家族中这一特殊的神经活动异常的表现形式在不同受累家族成员中是相当稳定的,这为患者提供了有用的预后信息。然而须牢记的是,很多常染色体显性遗传病如马凡氏综合征或神经纤维瘤等,具有很大的变异性。

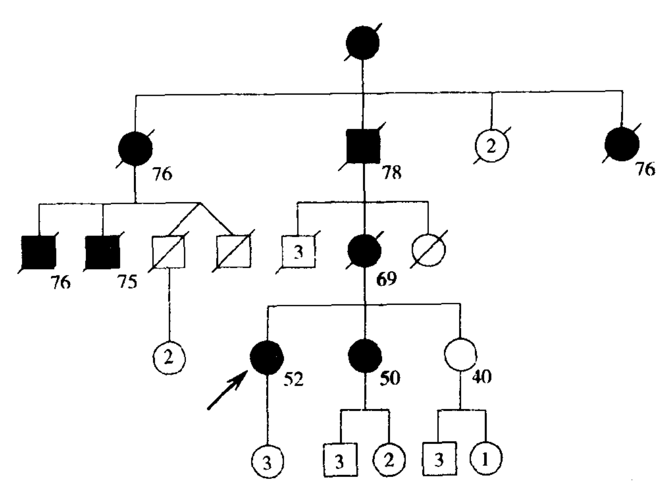


图 12-1 一个良性常染色体显性舞蹈病家系的系谱。患者用黑色表示。讨论见原文。

家族史的第三个也是最重要的应用是**遗传病的症状前诊断及临床疾病的预防或杜绝**。对家族史信息的正确解释可为具有患病风险的个体或家族提供遗传咨询,并进行遗传疾病的早期诊断和治疗。特别是,医生可以集中对小范围特定的高危人群进行干预。应该清楚的是,这种干预的机会也包括针对除患病个体外其他具有患病风险的家族成员的一种责任。以下几个例子将说明这些原则。

W.H.在诊断为家族性高胆固醇血症后,他所有生存的一级亲属被鉴定为具有这种相同疾病的高风险(50%)。尽管这种疾病常见(1/500),但 W.H.一级亲属发病率为一般人群的 250 倍。因为早期干预可以阻止动脉硬化的进展,所以受累者的子女是我们采取预防措施的重点。W.H.的三个孩子,尽管全部健康,但根据血脂检查患有家族性高脂血症,所有的三个孩 子在予以饮食和随后的药物治疗后在 2 岁、6 岁、9 岁时血清胆固醇水平

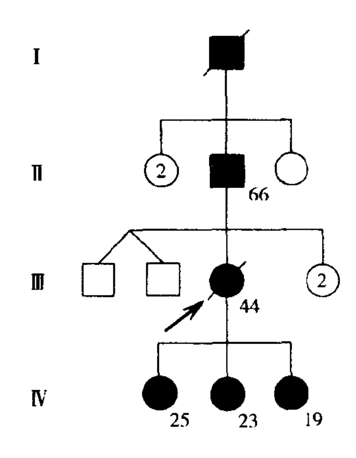


图 12-2 一个家族性结肠多发性息肉家系的系谱。讨论见正文。

正常。很有希望的是,这些早期干预可以减慢疾病的进程。

家族性结肠腺瘤样息肉(FAP)是应用家族史提供早期干预的另一个独特的例子(图 12-2)。先证者在她 42 岁时出现结肠大出血,并发现结肠多发息肉和结肠癌;18 个月后她死于骨转移。她的父亲因结肠癌而行切除术,活到 66 岁;他的父亲死于结肠癌。关注这一家族的外科医生对先证者的 3 个女儿的放射影像及内窥镜结果进行了研究。最大的女儿 18 岁时因发现结肠多发性息肉而行切除术。第二个女儿 16 岁检查时似乎正常,但 18 岁时再次复查发现多发性结肠息肉并行全切术。最小的女儿 16 岁时发病亦行结肠切除术。尽管在青春期切除结肠是相当大的手术,但这确实有效地防止了这些有患病风险的年轻女子结肠癌的发生。结肠多发性腺瘤基因(*APC*)定位克隆于染色体 5q 位置上,这为具有发病风险的家庭成员在症状出现前作出诊断提供了分子诊断学方法,讨论见第十一章。

家族性多发性息肉通常在 10 ~ 20 岁间作出诊断,而结肠或直肠癌多在 30 ~ 40 岁时发生,较一般人群发病早约 30 年(图 12-3)。然而,一些家庭显示早发复发现象,所以详细的家族史在疾病的预防上证明是很有用的。F.H.在 17 岁时诊断为多发性息肉,23 岁时发现患有结肠癌,转移后很快死亡。他的 3 个孩子中,2 个在 10 岁前有直肠出血,表明患有结肠多发性息肉。他们的初诊医生计划在 10 年后再评估他们的病情。然而,该家族中很早发生多发性息肉和先证者早发结肠癌,所以人们认识到这两个孩子会像他们死去的父亲那样在十多岁时就有发生结肠癌的高风险。因此,两人在 10 岁时都作了全结肠切除术,幸运的是,手术时未发现结肠癌。应用从家族史获得的信息可对这一特殊家族较其他一般的患家族性多发性息肉的家族进行更积极的治疗和更早期的干预。

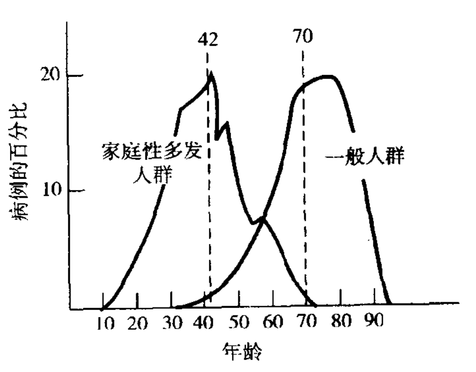


图 12-3 家族性结肠多发性息肉和一般人群中结肠癌患者的年龄分布(引自 Bussey HJR. Familial polyposis coli. Baltimore: Johns Hopkins University Press, 1975)。

图 12-4 描述的系谱表示应用家族史预防癌的发生。C.R.求助于医学遗传学诊所寻求 α_1 - 抗胰蛋白酶缺乏症的遗传咨询。然而,家族史也揭示她母亲在诊断为癌症后 11 年于 51 岁时死于乳腺癌。她母亲的祖母在 41 岁时诊断为乳腺癌,她母亲的侄女最近在 40 岁时也发现患有乳腺癌。这三个近亲在月经前发现乳腺癌,说明该家族乳腺癌患病风险至少高于一般人群的几倍(第十一章)。因此,对 C.R.来说建议比无家族史的年轻妇女进行更早、更经常、更彻底的随访检测乳腺癌是明智的。正如第十一章所讨论的,这种明显的家族史也可认为是可能发生了 *BRCA1* 和 *BRCA2* 突变的结果。

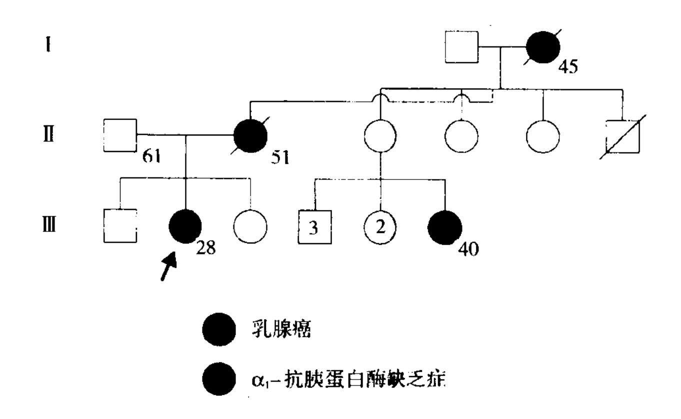


图 12-4 一个 α_1 - 抗胰蛋白酶缺乏症妇女的系谱。注意：先证者的母亲、外祖母、母方的第一代表兄妹（用黑色圆圈表示）有早发性的乳腺癌。

（一）如何获取家族史

表 12-1 列出了获得家族史的重点。家族史不必太长，但必须是经过认真推敲的。最重要的是，它必须与目前的临床疾病相关。尽管家族史通常包括各种各样的条目，包括“既往史”、“社会史”、“系统回顾”等，当它是“现病史”的一部分时才是最相关也是最有用的。这样，在获取家族史时应首先询问发生与引用的病例或患者相同或相关的疾病，而不是列出患有“家族倾向”的疾病。若患者疑有心脏病，则应询问其他家族成员是否患病或是否患有诸如糖尿病或高血压等与心脏病相关的疾病。常规询问过敏史、哮喘史、结核病史等任何一种家族史可能给医生一种病人患有一系列疾病的印象，或者更糟的是医生认为他或她已经获得了家族史，但这几乎不能提供有用的信息。除上述观点外，还要询问可预防或可治疗的疾病。因此，除询问与患者目前症状有关的疾病外，询问高血压、早发冠状动脉疾病史或多个家族成员的癌症早发病史（特别是乳腺癌或结肠癌），任何原因引起的早期死亡史都是很有价值的。通过这种方法，可以检测到常见病如高血压、冠状动脉疾病、癌症等高于平均水平发病风险的个体。

表 12-1 获取家族史

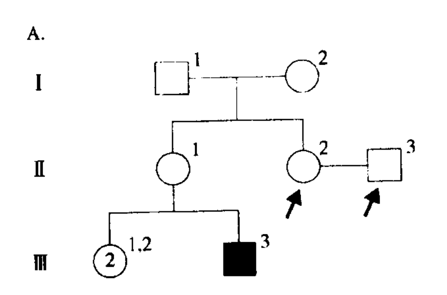
1.使它与“现病史”相关
2.询问早发性可预防的疾病
3.特殊询问所有一级亲属
4.询问可提供信息的亲属
5.记录种族和人种的背景
6.询问近亲婚配状况
7.保持家族史及时性

医生必须特别询问与索引病例有 50%相同基因的一级亲属情况(双亲、同胞、孩子),应该询问他们的年龄、健康状况,已经死亡的必须询问死亡的年龄和原因。还应特别询问婴儿死亡情况,因为有时这些信息是不报道的。在很多病例中,已确定的未受累亲属的信息可能是有用的,这将在下面讨论。当医生研究已知的遗传病时,询问有关亲属是极重要的。例如,当医生研究 X 连锁隐性遗传病时,应询问家族中母亲方面的男性亲属情况。相比之下,当研究常染色体隐性疾病时,应特别询问同胞的情况。而研究常染色体显性疾病时,应特别询问家族中受累一方所有一级亲属的情况。在获取这些信息的过程中,医生应对患病个体以外的人员表现出他(她)的责任。从这些其他亲属获得的信息可用来对他们进行早期干预或预防疾病的发生。

因为一些遗传疾病在不同种族、人种中发病率有明显不同(例如 Tay-Sachs 病、镰状细胞贫血、地中海贫血、囊性纤维化),记录家族中的种族背景是很有用的。最后,尽管近亲婚配在美国罕见,但还应询问疑患罕见常染色体隐性疾病个体的双亲近亲可能关系。

家族史经常是在急诊情况下获得的。然而,当这些信息在诊断中很有用时,也须牢记仅需花费一点时间就可询问到患者是否有相关疾病的家族史。家族史不是固定不变的。随着个体年龄增加,他们表现出晚发遗传病的机会就增加。具有儿童遗传病风险的家族孩子越多,可以获得的信息就越多。因此家族史信息必须是新的。

图 12-5 描述的系谱说明了可信的负面信息的重要性。一个年轻的医生和他的妻子(用箭头指出)曾寻求遗传咨询,因为妻子的姐姐(Ⅱ-1)有一个儿子死于 Duchenne 肌营养不良。图的上半部分(A)表示他们最初与我们诊所接触时提供的家族史信息。根据这一



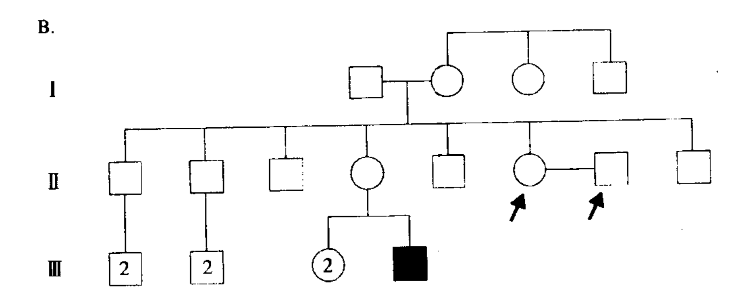


图 12-5　散发性 DMD 家系的系谱。图的上半部分(A)表示在最初与诊所接触时获取的信息。图的下半部分(B)表示进行遗传咨询时获取的家族史。注意:Ⅱ代中 5 个未受累的男性。咨询者用箭头指出。

信息,不能确定受累男孩的疾病是由于发生了新突变还是由携带者母亲遗传而来的结果。若是后一种情况,医生的妻子是携带者而生育患有 Duchenne 肌营养不良孩子的风险将增加。下半部分(B)表示这对夫妻进行遗传咨询后而有机会获取的更详细的家族史。从这个较大的家系可直观地看到妻子的母亲(Ⅰ-2)是 DMD 携带者的可能性更小,因为她的五个儿子都没有受累。因此,她两个女儿不可能是携带者,受累男孩的疾病更可能是发生了新突变的结果。用 Bayes 定理(下面讨论),这种"直观性"可被转换为用数学方法对寻求遗传咨询夫妇的发病风险作精确地评价。随着 *DMD* 基因的克隆(见第九章),可用分子诊断学方法检测 DMD 杂合性携带者。然而,由于无生存的受累男性或已确定的女性携带者,在这个家族中这样的试验是不清楚的;这种系谱信息在咨询中是很有帮助的。

正如上面所讨论的那样,医生的责任除涉及患者外至少应扩展到有遗传病风险的患者亲属,并对他们进行遗传咨询和早期干预。这已从上面讨论的家族性高胆固醇血症和家族性结肠息肉及图 12-6 所示的 Duchenne 肌营养不良的病例很好地说明了。前来遗传咨询的男子没有将疾病传递给其子女的风险,因为他 27 岁时仍然健康,因此他没有携带突变的等位基因。然而,家族史清楚地显示出他母亲的亲属、姨母、表妹、外甥女等都有发病风险,通过他得知必须对她们进行遗传咨询和携带者检测来提供信息。随着 DNA 诊断实验的应用,对这种疾病进行高度精确的携带者检测和产前诊断已是可行的。

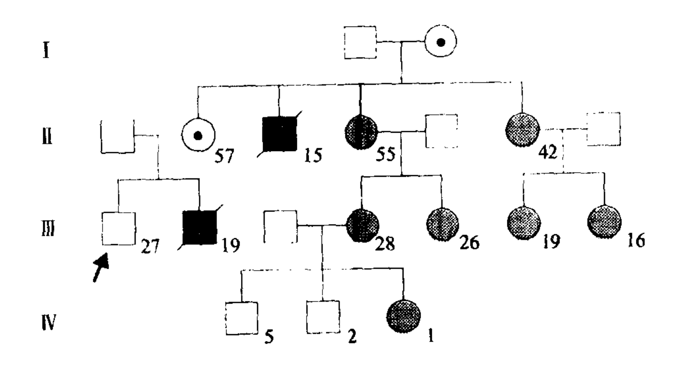


图 12-6 DMD 家系的系谱。注意有这种 X 连锁病携带者风险的女性亲属(用阴影表示)。

家族史也因各种非医学原因而获取,并以奇特的方式记录下来。图 12-7 表示 William Hogarth 的 6 个著名系列油画中的第一个"婚姻模式"。新娘和新郎的父亲正在谈论联姻的经济问题。一个父亲有一个以爵士祖先为核心的家谱画卷。假设,这个家系的质量是有价值的。图 12-8 描述英国伊丽莎白Ⅰ世家谱的一部分。这个 33 英寸的羊皮画卷把女王的祖先追溯到很早以前。这个系谱在伦敦北部伊丽莎白度过童年和青少年时期的地方 Hatfield Home 展出,以艺术家的手法绘出。最后,遗传学也和其他方面的科学及医学一样也会被魔鬼诅咒的。图 12-9 描述的两个系谱来自臭名昭著的 1935 年纽伦堡种族法所描绘纳粹观点,什么决定了犹太人和亚利安人的背景,仅在一些特殊环境下才允许或是严格禁止何种婚姻。



图 12-7 William Hogarth 画的"婚姻模式"。注意画卷右下角的系谱。这是挂在伦敦国家画廊的 6 个油画中的第一个。

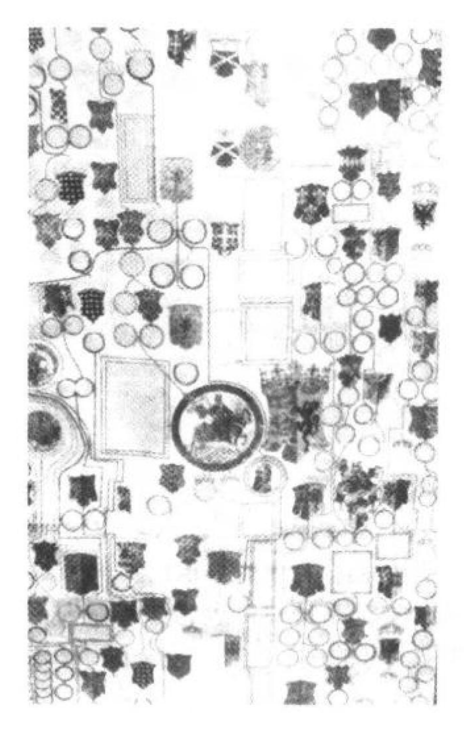


图 12-8 英国伊丽莎白Ⅰ世的系谱的一部分。中间重点指出的是征服者威廉，画卷在伦敦北郊的设菲尔德楼展出(引用获得允许)。

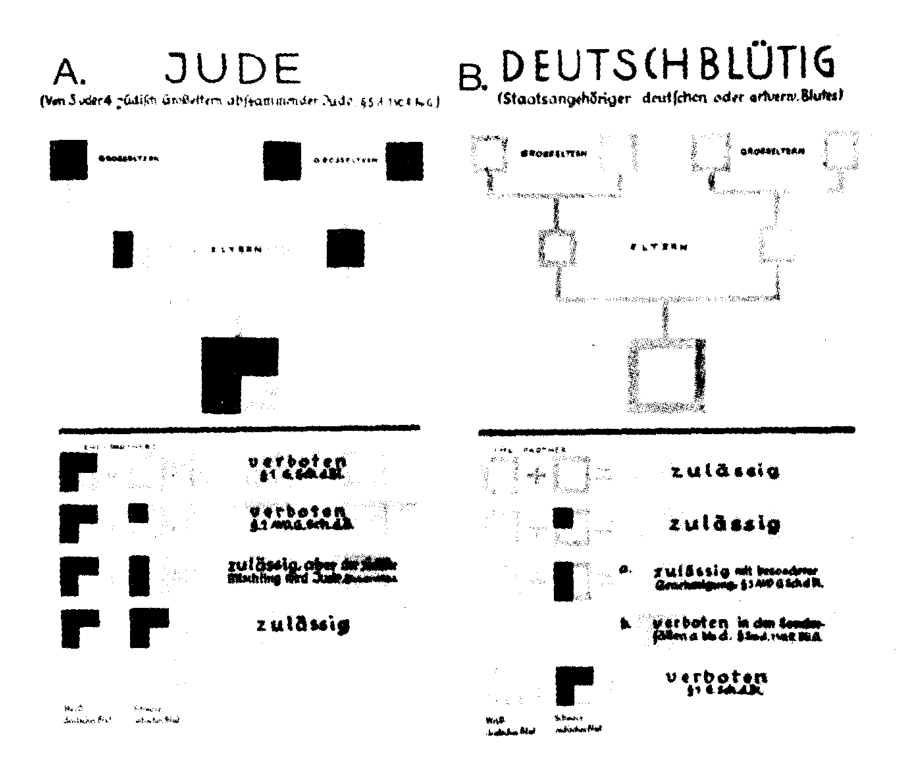


图 12-9 1935 年纽伦堡种族法确定的犹太人 A 和德国人 B 的系谱模式图(引用获得允许)。

(二) 医学遗传学中的体格检查

与其他医学领域一样,体格检查是评估遗传病患者的重要部分。尽管 DNA 突变分析技术发展迅速,很多遗传病如神经纤维瘤Ⅰ型和马凡综合征等疾病的诊断还基本以详细的体格检查为基础。体检在评价先天畸形儿童时是特别重要的,这些儿童在临床遗传学家所检查的病人中占有很大一部分。先天异常为儿童死亡和发病的重要原因,2% ~ 3% 的新生儿被诊断患有严重的先天异常,5%以上的 2 岁儿童发现患有异常。在美国,先天异常通常认为是多因素的,是出生后第一年死亡的最常见原因。

(三) 畸形学

畸形学(Dysmorphology)是临床遗传学领域中关于先天性解剖异常的诊断和治疗的部分。先天异常通常按定义分类如下。**畸形**(malformation)是确定为由于内部异常发育过程而引起的器官或部分器官的形态缺陷,如唇裂。相比之下,**破损**(disruption)是由于外部因素破坏或干预了原来正常发育过程而引起的形态缺陷。与畸形相比,受累器官发育潜能原来是正常的(图 12-10)。外在因素有创伤、感染、及药物引起的畸形。例如,桡骨发育不全或桡骨及相关结构正常发育受损可以是 Holt-Oram 综合征中心脏异常伴发的畸形。

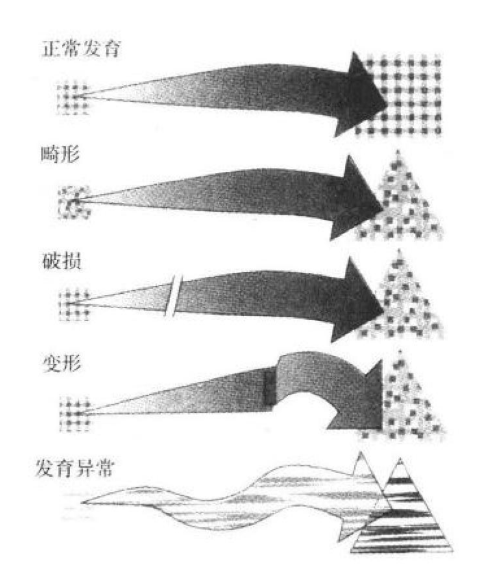


图 12-10 不同类型先天异常发育的代表模式图。异常原因、过程和结果图中指出。详见正文(引自 Spranger J, Benirschke K, Hall JG, et al. Errors of morphogensis: Concepts and terms. J.Pediatr. 1982; 100:160～165.)。

这是一种由编码转录因子的 *TBX5* 基因突变引起的常染色体显性遗传病。桡骨发育不全也可以作为破损而发生,是由药物反应引起的胚胎病的一部分。

变形(deformation)是由机械力引起的部分身体的形态、位置异常。这些机械力可以是作用于胎儿的外力,如羊水量的改变(见下文);或是内力作用的结果如神经、肌肉异常造成的胎儿活动减弱。例如:马蹄内翻足或棒节足可因羊水量减少(羊水量的缺乏)使胎儿受外力压迫引起,或因继发于脊髓脊膜突出的神经缺陷及先天性强直性肌萎缩造成的肌无力而引起的胎儿不运动。

发育异常(dysplasia)指细胞构成组织的结构和形态学异常。因为缺陷涉及到所有发现受累组织的解剖部位,发育异常分布范围较大。第三章讨论的软骨细胞发育异常引起的软骨发育异常是范例。

一个患者的多发性异常可呈因果或致病的关系。**顺序发生**(sequence)是一种由先前的单一异常或机械因素引起多发性异常的一种方式。Potter 的顺序发生由羊水量减少所致,而它是由继发于正常肾发育障碍或尿道阻塞的胎儿输尿管缩窄引起的胎儿尿量减少所致,或是羊水慢性丢失的结果。在任一情况下,羊水量减少导致胎儿受压,出现特征性的扁平脸、手足的位置及发育异常、肺低功导致呼吸不足和死亡。相比之下,**综合征**(syndrome)是多发性异常呈致病性相关的一种模式,并不表示简单的单个顺序。第三章讨论的马凡综合征就是熟知的一个例子。

临床上,重要的是判定一种先天异常是独立的异常或是组合方式的畸形,诸如顺序发生或综合征。例如,第四章讨论的一个婴儿患有唇裂而无其他畸形则预后良好,同胞的再发风险低。而另一方面如第八章所讨论的,唇裂只是 13 - 三体综合征一种特征,它是伴有多种其他威胁人生命异常的严重染色体疾病。

先天性异常的正确分类对诊断、处置和遗传咨询具有提示作用。畸形综合征和发育异常趋向于有确切的病因,在起源上是单基因病、多基因病、或染色体疾病。为提供确切的治疗及遗传咨询作出特异诊断是必要的。变形的再发风险依赖于机械因素。而破损是散发的,无明显再发风险。诊断先天性异常的有用信息可容易地利用市售的 CD-ROM 版本的两个研究性照片数据库来获得[POSSUM(Pictures of Standard Syndromes and Undiagnosed Malformations) Murdoch Inst. for Reaearch into Birth Defects. PO Box 1100 Parkville 3052, Melbourne, Australia, tel:(613)345 ～ 5045 ; and the London Dysmorphology Database. Oxford Univ. Press, 200 Madison Ave, New York 10016]。

二、遗传咨询

遗传咨询是一个交流过程，其目的是提供患有一种遗传病的个体和/或家族关于该病病情风险的信息，探讨此信息的个体后果，为有风险的夫妇提供作出有根据的生育决策的信息。下面遗传咨询的定义由美国人类遗传学会于1975年正式通过[From Ad Hoc Committee on Genetic Counseling (Epstein CJ Chainman). Genetic counseling. Am J Hum Genet 1975;27:240 ~ 242. Published by The University of Chicago]。

遗传咨询是一种交流过程，讨论人类疾病的发生、发病风险及在一个家族中发生的遗传病等有关问题。这一过程包括一个或几个训练有素的人员帮助个人或家族(a)了解医学事实，包括诊断、可能的病程、及治疗措施；(b)评价该疾病的遗传方式和在特定亲属中的再发风险；(c)了解处理再发风险的另外的方法；(d)采用与他们的患病风险、家族目标、情感和宗教标准及其决策相适应的行动方针；(e)对受累家族成员的疾病和/或该病的再发风险作出可能的最佳调整。

(一) 遗传咨询的指征

遗传咨询的指征见表12-2。尽管对已知或未知遗传病的患者或家族进行遗传咨询的价值已经在医学界得到广泛的共识，遗传咨询在其他方面的作用也同样的重要。出生缺陷和/或智力低下可发生于单基因病或染色体病，或是由于遗传因素与环境因素共同作用的结果。在每个病例中，确切的诊断性评估和遗传咨询是处置患者的重要部分并可使其他的家庭成员作出有根据的生育决策。35岁以上的妊娠妇女伴发Down综合征(见图8-18)和其他三体型的风险有所增加，应为她们提供遗传咨询和产前诊断。在多个家族成员中早发某些癌症常提示在其病因学中有强烈的遗传因素，这些因素易于用筛查试验检出。复发性流产有时与某个双亲的染色体平衡易位有关。某些病毒感染如风疹病毒(German measles)、某些药物如用于治疗囊性粉刺的维甲酸衍生物和用于治疗癫痫病的抗癫痫药物等已知都是致畸的。尽管要求不严格，但也须对在妊娠过程中接触潜在致畸物质的妇女提供遗传咨询。最后，正如第四章已讨论的那样，近亲婚配的后代出现罕见的

表12-2 遗传咨询的指征

1.在患者或家族中存在已知和可疑的遗传病
2.出生缺陷
3.智力低下
4.高龄孕妇
5.早发性癌症的家族史
6.习惯性流产
7.接触致畸因素
8.近亲婚配

突变等位基因纯合性的风险大大增加，因此易患常染色体隐性遗传病。这种风险的大小在不同民族、宗教、种族中有相当的变化。在理论上多基因病的风险也增加。应该注意的是在所有孕妇中，大约有3%的婴儿在出生时伴有严重的遗传病或出生缺陷。经验表明，一级堂兄妹婚配其后代发病风险高达2~3倍。

(二) 遗传咨询传达的信息

表12-3列出了我们认为在遗传咨询中需要提供的几种主要信息。每一种都将讨论如下。

表12-3 遗传咨询中所传递的信息

1. 风险和再发风险的幅度
2. 疾病对患者和家庭的影响
3. 减轻影响或风险的可能性
4. 预测未来的发展

1. 发病风险或再发风险的幅度

为提供确切的遗传咨询，必须对有风险的夫妇或家族的疾病作出正确诊断。所以，随时可能的检查受累患者，获得适当的有帮助的实验分析和所有相关的医学记录是主要的。第三章中讨论的遗传异质性在精确遗传诊断过程中会产生重要问题。图12-11所示为两个短肢侏儒外表相似婴儿的照片。左边的孩子患有软骨发育不全，以常染色体显性

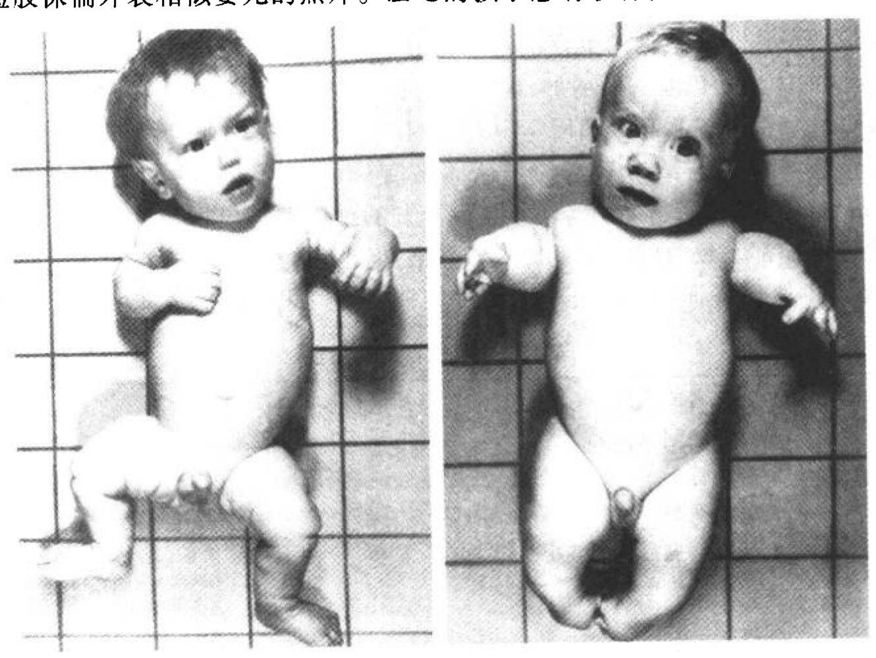


图12-11 两个短肢畸形的婴儿。左侧为软骨发育不全。注意大头、汤勺样鼻梁、肢体近端短缩。右侧为弯曲变形的侏儒。注意异常的拇指和螃蟹样脚(引用获得允许)。

方式遗传,通常是发生新突变的结果,而其双亲完全正常。在这种情况下,对此双亲来说,再发风险为零。相比之下,右边的孩子为侏儒畸形,以常染色体隐性方式遗传。因此,此孩子表型正常的双亲每再生一个孩子患这类侏儒的再发风险为25%。在这样的病例中不确切的诊断将导致不确切的灾难性的咨询后果。

表现度不一致在诊断上也会产生问题。在神经纤维瘤病病例中,在某些个体中表现型是适度的,他们可能是正常的,而某些受累儿童则认为是可能发生了新突变。在神经纤维瘤中,大约一半的先证者其双亲表型正常;他们的疾病认为是发生了新突变。所以,必须对他们的双亲进行彻底的检查,以确定其中的一个或另一个没有受累。如果双亲之一受累,其每个再生子女将有50%的患病机率,而不是第二个新突变的相当低的发病风险。

通过详细的系谱分析通常可以确定一种疾病的遗传方式。如果家族史不能提供信息但又必须作出确切的诊断时,Mckusick的人类孟德尔式遗传将是极有帮助的。这一目册提供了5000多种孟德尔式性状的遗传方式的信息,以及近4000种疾病的有用的临床信息。这些资源定期更新,且可通过网上(OMIM)与有关疑问疾病的大量参考文献直接相连的有价值的链接点而获得。在国际互联网上网址为:http//www.ncbi.nlm.nih.gov/omim。

Bayes定理

为提供最佳的再发风险估计,特别是当某一咨询者的基因型尚不清楚时,利用所有可以获得的信息是很重要的。这种情况通常利用Bayes定理(以18世纪英国科学家——一位牧师的名字命名)来解决。在描述Bayes定理之前,在此介绍一些概率的基本原则是适当的。

相加性原则(additivity principle)说明如果两个事件(A和B)是彼此独立的,其中一个或另一个概率为两个事件各自概率之和。

$$\text{Pr}(\text{A 或 B}) = \text{Pr}(\text{A}) + \text{Pr}(\text{B})$$

例如,在一个(或一对)骰子中一次抛出一个2或一个3的概率是1/6+1/6,或1/3。

独立性原则(independence principle)说明两个或多个独立事件联合发生的概率是它们各自概率的乘积。

$$\text{Pr}(\text{A 和 B}) = \text{Pr}(\text{A}) \cdot \text{Pr}(\text{B})$$

例如,一次抛出两个骰子同时获得同一个面的概率为 $1/6 \times 1/6$,或1/36。

两个不相关事件联合发生的概率是第一个事件发生的概率乘以假设发生第一事件后发生第二事件的概率的乘积。后者称作**条件概率**(conditional probability).注意:Pr(B/A)表示既定A事件后B发生的概率。

$$\text{Pr}(\text{A 和 B}) = \text{Pr}(\text{A}) \cdot \text{Pr}(\text{B} | \text{A})$$

玩牌时,从52张牌中抽出黑桃A的概率是1/52,但是如果你抽出的牌已经是黑桃,那么你抽出黑桃A的概率是1/13。

Bayes定理表示假设已经发生某一事件(E),一组可能引起该事件发生的所有原因(C_i),那么某一特定原因C_i引起事件E发生的概率[写作$Pr(C_i/E)$]等于该原因出现的概率$Pr(C_i)$,乘以该事件对于该原因的概率$Pr(E|C_i)$,除以事件E发生的总概率。事件E发生的概率等于所有引起事件E发生的潜在因素的概率乘以事件E对于每个因素的概率的总和。可写为等式:

$$Pr(C_i|E)=\frac{Pr(C_i)Pr(E|C_i)}{Pr(E)}$$

$$Pr(E)=\sum_i Pr(C_i)Pr(E|C_i)$$

Bayes定理常用于医学遗传学,E常指某一疾病的受累或未受累的个体,或有某实验结果的人,C_i表示特定的基因型。

Bayes定理的应用参见表12-4中的一个简单例子。囊性纤维化是一种常染色体隐性遗传病,患儿未受累兄弟是该病携带者的概率是多少?在这一实例中,事件E指临床上未受累的兄弟,原因C_i指他对突变的*CFTR*等位基因杂合性。如图3-14所示为常染色体隐性性状的婚配图解。理论上讲,当他的表型未知时,兄弟是携带者的概率为1/2,称做**前概率**(prior probability)。假设他是携带者,临床上表型正常的概率等于1,这叫做**条件概率**(conditional probability)。他是携带者并且临床上正常的**联合概率**(joint probability)为前概率与条件概率的乘积,或1/2×1=1/2。这是上述等式的分子。等式的分母是他在临床上的不同正常形式的概率之和,或基本上是不同的联合概率之和。例如,限定在分子之中的这些形式的第一种可能是:他是一个携带者。第二种可能性是:他是正常的纯合性个体,前概率为1/4。如果他是纯合性正常,他临床上正常的条件概率当然是1。他的基因型为正常纯合性且临床正常的联合概率是1/4×1=1/4。这样他是临床上正常的联合概率总和为1/2+1/4或3/4。而现在他是临床上正常的携带者,其**后概率**(posterior probability)或相关似然性$Pr(C_i|E)$是1/2÷3/4或2/3。这一结论如图12-12的饼形图所示。"饼形图"中临床表型正常相应地可分为三部分;其中两部分代表杂合性携带者的基因型Aa。总之,未受累兄弟是携带者的前概率为1/2。而考虑到他是临床正常的其他信息,估计他是携带者的概率可限定为2/3。在这个简单的例子中,上述限定值没有很大区别;而在下面的例子(图12-13、12-14和表12-5)中Bayes定理应用的重要性更加明显。

表12-4 常染色体隐性病中携带者状况的贝斯计算

	未受累同胞是杂合性携带者	未受累同胞是纯合性正常者
前概率	1/2	1/4
条件概率	1	1
联合概率	1/2×1=1/2	1/4×1=1/4
后概率	$\frac{1/2}{1/2+1/4}=2/3$	$\frac{1/4}{1/4+1/2}=1/3$

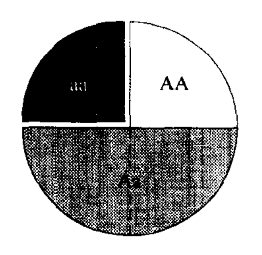


图 12-12　饼形图表示临床正常同胞是常染色体隐性性状携带者(Aa)的概率为 2/4 (db)3/4 或 2/3。

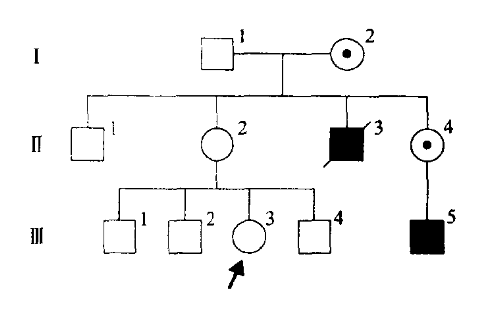


图 12-13　DMD 的家系系谱。讨论见正文。

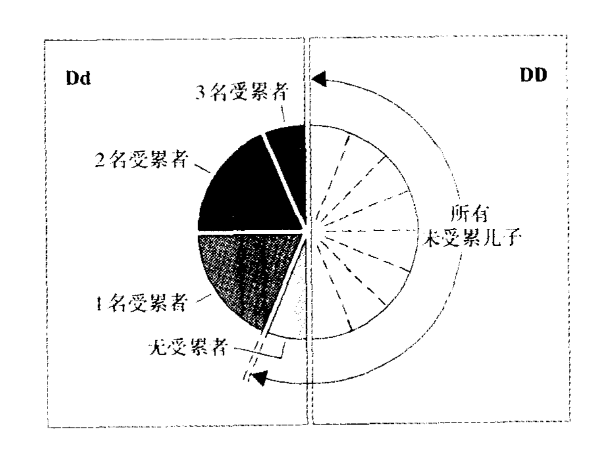


图 12-14　饼形图表示图 12-13 系谱中Ⅱ-2 在假设她有 3 个正常儿子的情况下是 DMD 携带者的概率。纯合正常人(DD)的前概率用不全的半圆表示。携带者(Dd)的前概率用圆的黑色部分表示。从两面扩展,若Ⅱ2 是携带者,可计算出她有 3 个患儿的概率为 1/8,有 2 个患儿的概率为 3/8,有 1 个患儿的概率为 3/8,无患儿的概率为 1/8(浅灰色部分)。在第 9 部分代表所有正常儿子,浅灰色部分代表 1/9。Ⅱ2 有 3 个患儿并是携带者(Dd)的概率为 1/9。

表 12-5　X 连锁隐性遗传病携带者 Bayes 计算

	Ⅱ-2 杂合性携带者	Ⅱ-2 纯合性正常者
前概率	1/2	1/4
条件概率	$(1/2)^3 = 1/8$	$(1)^3 = 1$
联合概率	$1/2 \times 1/8 = 1/16$	$1/2 \times 1 = 1/2$
后概率	$\frac{1/16}{1/16 + 1/2} = 1/9$	$\frac{1/2}{1/2 + 1/16} = 8/9$

图 12-13 所示为 Duchenne 肌营养不良的家系。咨询者Ⅲ-3 其基因型未知,想咨询她是否是 Duchenne 肌营养不良的携带者。如果她是携带者,她一定遗传了她母亲的突变等位基因。所以,首先要回答的问题是她母亲Ⅱ-2 是否为携带者。根据系谱提供的信息可

得出结论Ⅰ-2和Ⅱ-4是Duchenne肌营养不良携带者。Ⅰ-2有一个受累儿子和一个受累孙子，Ⅱ-4有一个受累兄弟和一个受累儿子。Ⅱ-2是携带者的前概率是1/2。而另外的信息是有用的(已发生的时间E)，即她有三个未受累的儿子。如果Ⅱ-2是携带者，她任何一个儿子受累(或未受累)Duchenne肌营养不良的条件概率是1/2。而她三个儿子都未患病的概率是这几个独立事件概率之积，$(1/2)^3$或1/8。她是携带者且她的三个儿子未患病的联合概率是$1/2\times1/8$或1/16。如果她不是携带者，例如是正常的纯合子，她的三个正常儿子的条件概率当然是1，她是表型正常的纯合性和有三个正常儿子的联合概率是$1/2\times1^3=1/2$。因此Ⅱ-2有三个正常儿子且是携带者的后概率是

$$\frac{(1/2)(1/8)}{(1/2)(1/8)+(1/2)(1)}=\frac{1/16}{1/16+8/16}=\frac{1}{9}$$

考虑到她有三个正常的儿子，Ⅱ-2是Duchenne肌营养不良携带者的概率由1/2降到1/9。Ⅲ-2是Duchenne肌营养不良携带者的概率是其母亲的1/2，即1/18，大约为5%。不利用她三个未患病兄弟的信息，她的概率是1/4或25%。在这一家族中，Bayes定理的分析还可以用分子诊断试验来加以补充，以便明确地诊断Ⅲ-2是否是Duchenne肌营养不良的携带者。

Bayes定理被应用于临床遗传学的各种情况中。在上述讨论的例子中，它可限定遗传咨询中再发风险的预测。Bayes定理也被应用于解释遗传筛查及诊断实验结果以确诊所有疾病。因此，努力理解Bayes定理的原则和对它的恰当运用是医学实践的重要部分。

2. 疾病对患者及家庭的影响

人们常有一个错误的理解，认为对遗传病的再发风险的估计是遗传咨询的全部，而其实这仅是第一步。不了解疾病的影响，而估计其再发风险大小几乎是没有意义的。疾病的影响指对患者和/或家庭造成的临床、经济、社会影响和思想负担。对遗传病影响的咨询必须包括对其自然病史、临床处置、疾病的预后、及心理、经济和保险认定等的公开讨论。

例如，短指趾症(图12-15)包括短指和短趾的一组特征的病况。这种疾病以常染色体显性方式遗传，所以每一个受累个体的后代大约有50%的再发风险。而该病仅对容貌有影响，受累个体仅有轻微的功能障碍。相比之下，对于生有21三体综合征患儿的妇女再发风险为1%，风险很低。然而在引起智力低下和严重出生缺陷方面，Down综合征的影响是巨大的。因此，再发风险的大小一定要在疾病的影响方面来解释。


图12-15　短指(趾)症患者的手。

任何一种疾病随其自身严重程度、受累个体和受累家族的不同，其影响也不同。强调这一点是很重要的。Down综合征是一个很好的例子。在一些家庭中抚养这样的孩子可以丰富他们的人生经历，但在另一些家庭中却是毁灭性的。遗传咨询家必须牢记对于每个家庭来说，评价某一疾病的影响应该是非常不同的。

3. 改变疾病的影响和/或发病风险

对于一些遗传疾病，现在通过特殊治疗和遗传实验来改变它的影响和再发风险是可能的。如果疾病是可治疗或可预防的，如继发于家族性结肠息肉的结肠癌，或继发于苯丙酮尿症的智力低下(下面讨论)，疾病的影响即使重大，也与不能治疗和预防的疾病有很大不同。如果一些疾病可以在产前作出诊断，如 Down 综合征及镰状细胞贫血等疾病，不管胎儿是否受累，遗传学家们也可在一定程度上改变此病的再发风险。最后，人们的支持群体、教育和社会背景、生育观点、辅助性生育技术的应用等都将影响家族对疾病再发风险和影响的看法，也将影响他们的生育决策。

现在有关临床遗传学的重要信息可在 internet 网上获得，包括辅助诊断、实验检测、临床遗传咨询及对患者和家庭提供的信息和帮助性服务等。获得遗传专业信息的有效途径在 Kansas 大学医学中心主页上，包括各种临床和科研服务的链接点。网址：http://www kumc.edu /gec/ geneinfo.html。密歇根大学对患者及专业人员提供的服务可通过密歇根遗传服务主页获得，网址：http://www.mdmh.state.mi.us/pha/bcfs/cshcs/genetics/。还可以通过密歇根大学人类遗传系主页，网址：http://www.med.umich.edu/hg/home.html 或遗传学会主页网址：http://www.faseb.org / genetics/ 获得其他信息。后者还包括遗传培训、证明和职业等有用的信息。

4. 未来发展的展望

医学遗传学是一个飞速发展的领域，几年前还不能做产前诊断和治疗的疾病现在却可以了。所以咨询家应当指出在不远的将来能够进行有用的最佳预测。人们常在她们生育年龄的早期就进行遗传咨询。因此如果他们有二十年的时间去决定是否要孩子，了解在这一段时间内情况是否会发生重大变化是非常重要的。如在本章后面讨论的那样，一些遗传疾病的产前诊断已有飞速的发展。

所有这些问题将在下面一封三十年前可能患有 Huntington 病的年轻妇女来信中反映出来。

尊敬的阁下：

我写这封信想向你咨询遗传病——Huntington 舞蹈病。我是一名 18 岁的大学生。我的母亲、外祖父和曾祖父都患有此病。

我打算在三年内结婚。我的男友知道我母亲的病情，但我还没有告诉他遗传的所有状况。我想知道节育是否是惟一的选择，或者还有任何新进展。人工放置供体的卵子到不孕妇女体内的技术是否已有所发展？

我知道我母亲的病情已影响到父母的婚姻，我应该彻底(原文如此)放弃结婚的希望吗？或者在我的孩子长大的时候，Huntington 病的治疗已得到发展，这种想法是错误的吗？

请坦率地告诉我，因为我有足够的勇气面对事实并会采取最佳的措施。感谢您为我花费的时间及精力。

1969 年 12 月 5 日

这位非凡的妇女提出了几个遗传咨询中的重要问题。她知道自己家系的遗传方式及再发风险的幅度。因为她与患病的母亲生活在一起,充分了解此病不仅影响到其父母的婚姻,也影响到患者本人及家庭。她想询问是否放弃生育自己的孩子和收养孩子是其惟一的选择,或者还有其他的可能性(体外受精和移植供体者的卵子)。她也想问是否未来的发展可以改变下一代的病情。她提出的问题有惊人的洞察力和预见性。信写于1969年,当时将他人的卵子植入不孕者体内简直是一种幻想。到1997年,尽管费用还很高,但已建立了体外受精及将他人卵子植入不孕者体内的生育方法。药物疗法尽管还不能完全有效的治愈Huntington病的运动障碍,但此症状已得到改善,对痴呆却没有治愈及有效的方法。然而对Huntington病分子遗传机制的认识有了显著的进展,使人们可以对此病直接进行症状前和产前诊断的试验。因此,提供给这位年轻女子的生育决策,现在与她写这封信的时候相比已有很大的区别。

(三)遗传咨询的过程

遗传咨询的第一步是确定咨询者实际想咨询的问题。我们认为在遗传咨询中必须提供的主要信息已在前面讨论过了。但事实上咨询者可能寻求非常具体的服务,包括咨询如何对待内疚和焦虑感。为了最好的帮助遗传咨询者,咨询家必须仔细地确定他想咨询什么。由于遗传咨询是一种交流过程,必须考虑到信息传递的各方面。必须考虑到咨询者情感和智力等方面的接受能力。一些夫妇指出当他们在一起或与他们的亲戚、朋友在一起接受遗传咨询时,他们能更好地理解和接受咨询。应该给咨询者传递多少信息及在什么水平?显然,必须用咨询者能理解的语言进行信息的传递。我们确信坦诚是传递这样信息的最佳方式。当人们在产生焦躁与提供非真实性良好预后间必须作出权衡时,困难问题便产生了。神经纤维瘤病再次提供了一个很好的例子。人们不能预测哪一种并发症会发生在一个受累个体身上,更不能对其做成功的治疗,告诉咨询者所有可能出现的并发症是否合乎常理?另一方面如果不这样做可能会有不切实际的期望,依据此信息做出的生育决策是无保障的,患者对他或她将要遇到的潜在医疗问题没有警惕性。

信息的交流有重要的个人推论,经常包括对一些个体来说是较难理解的医学或科学的信息。重要的是咨询以一种适合患者的情况和理解的水平来传达。口头咨询也必须附有一份用患者可以理解的书面语言书写的函件,使咨询者与医生、家庭成员、牧师及其他帮助者反复阅读,给他们一段时间来接受这一信息。我们相信遗传咨询不是指令性的。信息应以非决定性方式来传达,以鼓励咨询者作出他或她的决定。生育决策是高度自主性的并应由本人而不是由医生和咨询家来作出决定。

探讨心理问题是遗传咨询的主要部分。有先天缺陷和/或遗传病患儿的双亲常常感到内疚、焦虑、挫折、愤怒。咨询家们必须协调这些问题且应具有支持性。他(她)也必须为咨询者提供有关社会和其他援助的信息,对这样家庭来说可能是非常有帮助的。在许多遗传病诊所,由一些卫生专业人士来进行遗传咨询,包括医生、遗传学家及由美国遗传咨询委员会授权的经特殊培训的高级遗传咨询家。咨询家们利用社会心理学方法对病人进行耐心的教育和咨询。利用这种方法使遗传学家们得以深入了解婚姻关系及人际

关系,得知患者希望解决的问题,包括夫妻双方的讨论和决定,了解患者的信仰、价值观、自信心等背景资料,并了解以前的健康状况、伤残、风险及死亡情况等。咨询家可预见常见的反应及出现的问题,查明患者的配合情况,研究在遗传咨询时交流的信息对病人的影响,调动患者的自主性及心理适应能力,发现帮助患者的群体,当他们需要附加随访治疗时对患者进行分类。

三、遗传筛查

遗传筛查是常规医疗监护的重要部分,一般针对可治疗疾病的早期诊断而言。我们熟悉的例子包括检查子宫颈癌的巴帕尼科拉乌实验(PaP)、检查早期乳腺癌的早期胸部X射线透视法、检查高血压的血压筛查、检查结核的皮肤试验。有时筛查与诊断的试验是相同的;而更常见的是筛查用于在群体中检查出一批需要进行特异诊断的人群。有效的筛查试验应该是廉价的,可应用于大群体,并能有效的检出需要进行特异诊断的一批群体。因为运用于医学遗传学实践,筛查试验主要有两类。第一种是遗传病受累个体的早期筛查,第二种是传递遗传病风险个体的鉴定。

因为任何筛查试验结果可以是阳性或阴性的,接受筛查的个体在筛查试验中可以是受累(或携带者)者或未受累者,试验可能有四种结果(表12-6):A受累个体试验结果阳性;B未受累个体结果阳性;C受累个体结果阴性;D未受累个体结果阴性。

表12-6 筛查结果

	受累者	未受累者
试验阳性结果	A	B
试验阴性结果	C	D

实验的**敏感性**(sensitivity)是疾病存在时试验阳性结果的百分率,或者是试验阳性结果受累个体的比例[A/(A+C)]。而**特异性**(specificity)指未发病时试验阴性结果的百分率,或试验阴性结果未受累个体的比例[D/(B+D)]。**假阳性率**(false-positive rate)[B/(A+B)]为未受累个体试验结果阳性的比例,**假阴性率**[C/(C+D)](false-negative rate)为受累个体试验结果阴性的比例。因此,试验的敏感性反映的是该试验能检测出受累个体的能力,而特异性表示的是该试验所排除未受累个体的能力。筛查的有用性——鉴别受累及未受累个体的能力不仅依赖于阳性及阴性结果的区别能力,也依赖于试验的**阳性预测值**(positive predictive value,PPV)[A/(A+B)]。PPV与待测群体的发病率密切相关,这一点可用例子说明。

让我们假设开展一个试验以检测罕见的常染色体隐性疾病的杂合性携带者,这个携带者筛查试验的敏感性为95%,特异性为95%。该试验用于诊断100例诊断患有D病的患者临床上未受累的同胞,获得了试验结果(表12-7)。回顾一下,未受累同胞是携带者的概率为2/3,这样就推测有67个携带者(Dd)和33个纯合性正个体(DD)。假阳性率为2/66或3%,而假阴性率为3/34或9%。该试验PPV为97%,显示可高效率地鉴别携带者及纯合性正常同胞。

表 12-7　在未受累同胞中筛查杂合性携带者

	携带者(Dd)	正常者(DD)
阳性结果	64	2
阴性结果	3	31
总　　计	67	33

但是现在考虑将同样试验用于 Ann Arbor 人群中，密歇根有大约 100 000 人参与试验。假设疾病 D 发病率为 1/40 000 或 0.000025。假设 Hardy-Weinberg 平衡，携带者的频率为 $2pq=0.01$。同样筛查试验应用于 Ann Arbor 人群的结果见表 12-8。在这些条件下，假阴性率是 50/94 000 或 0.05%。而假阳性率是 4 950/5 900 或 84%。换言之，假阳性结果的数量比真正的阳性结果高出 5～1 倍，而且经过这一试验检出的大多数"携带者"被错误分类。该人群体 PPV 值为 16%。因此，在两组人口数量相近的家族中，可高效从纯合性正常个体中鉴别携带者的试验，当纯合性正常个体数量远远超过携带者的一般人群时却出现了严重的错误。然而这一试验也是很有用的，因为阳性试验可从一批总群体(5 900/100 000 的个体)中比一般人群(1/100 或 1%)能鉴别携带者很高的概率(950/5 900 或 16%)。为高度精确地鉴别携带者，应对筛查结果阳性者进行特异的诊断试验。

表 12-8　在一般人群中筛查杂合性携带者

	携带者(Dd)	正常者(DD)
阳性结果	950	4950
阴性结果	50	94 050
总　　计	1000	99 000

遗传筛查试验的类型

遗传筛查有两大类型。第一种目的是早期检出受累个体，而对这些人进行医疗干预时会对患者个人和/或家族成员产生有益的效果。胎儿筛查是其中之一，包括对疾病如 Down 综合征的产前诊断；新生儿筛查的典型例子是对苯丙酮尿症(PKU)的筛查；及现在对成年发病潜能的疾病如结肠癌和乳腺癌的筛查。产前诊断的特殊形式是筛查妊娠神经管缺损或 Down 综合征胎儿的孕妇母体中的血清甲胎蛋白。遗传筛查的第二种主要形式是鉴别传递遗传病风险的个体。这种筛查试验的经典例子是 Tay-Sachs 病、镰状细胞贫血和地中海贫血等疾病的携带者检测实验。

遗传病的产前诊断将在本章后面讨论。一种间接的胎儿筛查是母体血清甲胎蛋白(maternal serum α-fetoprotein，MSAFP)的检测。有研究表明怀有开放性神经管缺损(脊髓脊膜突出)或开放性腹壁缺损胎儿的孕妇血清中甲胎蛋白(AFP)水平增高，这反映了羊水中 AFP 水平增高。由于血清中 AFP 水平可随妊娠时间和胎儿数量的改变而有相当大的变化，所以利用对胎儿作超声检查来精确地确定妊娠时间是很重要的。现在已知妊娠 16～18 周时两次连续的 AFP 水平增高提示孕妇怀有开放性神经管缺损胎儿的风险大约为

1/20。与此相比，在一般人群中的发病率约为1/1 000。筛查试验显示这群妇女有高于正常人群50倍的发病风险，因此在对她们进行神经管缺损产前诊断时须运用高分辨率的靶向超声检查和/或利用羊膜穿刺术检查羊水中AFP值。估计利用MSAFP增高能检测80%～85%的开放性神经管缺损。

在进行这种筛查项目过程中，发现低水平的MSAFP与Down综合征及其他三体型的高发风险有关，运用羊膜穿刺术在35岁以下的年轻妇女中可检出15%～20%Down综合征出生婴儿。在母体血清中同时检测MSAFP与母体血清中绒毛膜促性腺激素(HCG)和雌三醇("多重标记筛查")时，在35岁以下的妇女中可检出50%～60%Down综合征出生婴儿。由于对35岁以下的妇女一般不作常规的产前诊断，而现在75%的Down综合征出生婴儿发生在这一年龄组妇女身上，因此对于Down综合征来说，多重标记筛查是一种重要的遗传筛查试验。美国医学遗传学学院最近提出要对所有35岁以下的孕妇作遗传筛查。

(一) 对于可治疗和/或可预防疾病——苯丙酮尿症的新生儿筛查

新生儿筛查最佳示例是检测苯丙酮尿症(phenylketonuria, PKU)的Guthrie实验。PKU是常染色体隐性遗传病，发生在北欧人群中，在活婴中的发病率为1/10 000。如果不治疗，其特征为小头畸形和智力低下。此病由肝脏中可将苯丙氨酸转化为酪氨酸的苯丙氨酸羟化酶缺乏引起，其特征为血浆中酪氨酸浓度降低，而苯丙氨酸浓度升高，排泄含苯丙氨酸代谢产物——如丙酮的"鼠臭"味尿液。如果在新生儿早期严格限制饮食中苯丙氨酸摄入，可降低血浆中苯丙氨酸水平，防止出现明显的智力低下。

Guthrie实验基于观察β-2-噻吩丙氨酸可抑制枯草杆菌的生长，加入苯丙氨酸后共抑制作用被解除(图12-16)。从新生儿(出生了3天)身上采集一滴血滴在一张滤纸上干燥。将含有干血斑的纸片放在含有噻吩丙氨酸和枯草杆菌的琼脂上。如不含苯丙氨酸，细菌的生长将被噻吩丙氨酸抑制；在纸片周围有细菌生长，说明血样中存在足够的苯丙氨酸来解除抑制。该试验可用含有已知量苯丙氨酸的纸片来定量；通常苯丙氨酸浓度在4mg/dl以上时出现阳性结果。每一个阳性结果必须随后测量血浆中苯丙氨酸及酪氨酸的浓度来确定PKU的诊断。

每20个Guthrie试验阳性结果中，只有一个新生儿患有苯丙氨酸水平超过20mg/dl(正常值低于2mg/dl)、血浆中酪氨酸低水平的典型PKU。这些孩子必须尽可能地长期严格限制苯丙氨酸的摄入以预防智力低下的发生。一少部分孩子有持续、中等程度的血浆苯丙氨酸水平升高(通常低于10mg/dl)，以血中高苯丙氨酸为特征，现在已知此病继发于苯丙氨酸羟化酶等位基因的再次突变，导致这种酶的中度缺乏。这些孩子不需要严格限制饮食，而大多数也不会发生智力低下。Guthrie试验出现假阳性结果的最常见原因是由于肝脏发育不成熟导致暂时性高酪氨酸血症。这些孩子血浆中丙酮及苯丙氨酸水平都很高，而随时间推移二者水平下降，无须进行任何特异治疗。最后，PKU的新生儿筛查已经检测出不能用限制苯丙氨酸摄入方法来预防的以血中高水平苯丙氨酸和严重智力低下为特征的稀有病例。这些婴儿有嘌呤合成异常，导致苯丙氨酸羟化酶一个重要的辅助因子缺陷。

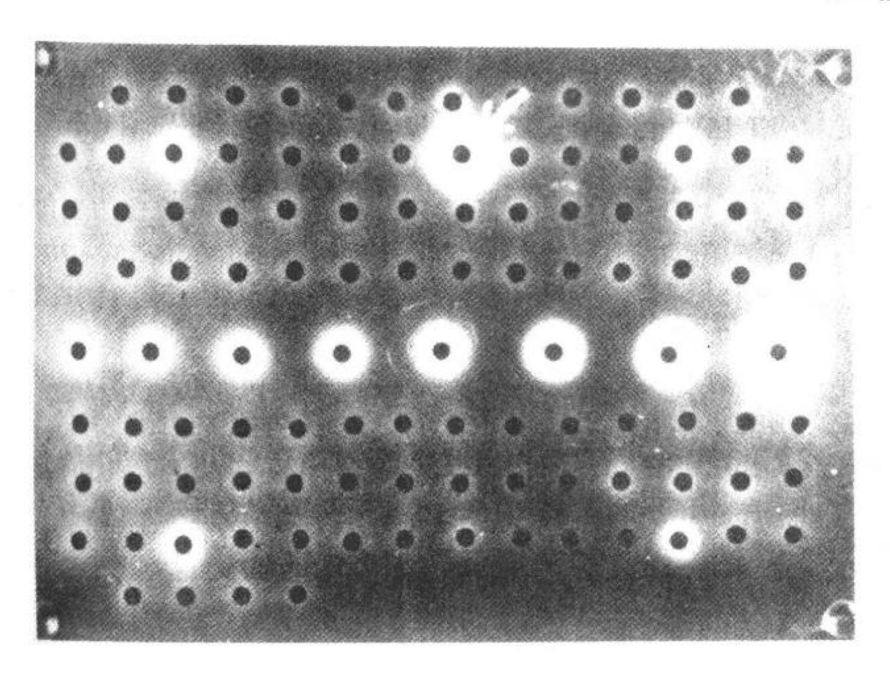


图 12-16　细菌抑制试验或 Guthrie 试验。第 5 排表示对照组苯丙氨酸浓度从 2mg/dl 到 5mg/dl。箭头指出第 2 排的一个阳性试验。

在 20 世纪 60 年代，PKU 新生儿筛查广泛应用以前，估计由 PKU 导致的智力低下约占 1%。Guthrie 试验在新生儿中的广泛应用和早期的饮食疗法已经使 PKU 不再是产生智力低下的主要原因。因为存在虽很低但确切的假阴性率，特别是在出生后的三天内血中苯丙氨酸可能增高（胎儿由于母体的新陈代谢而避免出现了血中高水平的苯丙氨酸）的情况下就做早期检测，一些受累婴儿就被漏诊。

遗传筛查项目取得了显著成功，也产生了难以预料的新问题。患 PKU 的妇女在儿童期成功地采用饮食疗法获得成功，但在成年后不再限制苯丙氨酸摄入，已发现生育了患小头畸形和严重智力低下的孩子。虽然所有这些孩子至少肯定是 PKU 杂合子，但仅有 1% 可患 PKU。智力低下推测是在子宫中暴露于母体循环中高水平苯丙氨酸的结果。为防止这种形式的智力低下发生，通过终生持续限制饮食和在整个妊娠过程中将苯丙氨酸水平保持在 10mg/dl 以下是有可能的。

一旦运用筛查项目作 PKU 筛查时，如果在干血斑上同时加入其他的新生儿筛查就会降低成本。例如，密歇根对所有新生儿（每年约有 135 000 人）筛查 PKU（一种快速流式荧光检测替代 Guthrie 试验）、先天性高酪氨酸症（一种伴发生长及智力低下的散发疾病）、镰状细胞贫血、先天性肾上腺增生（一种类固醇激素合成紊乱，伴随女性胎儿男性化及可引起休克和死亡的电解质紊乱）、半乳糖血症（一种先天性的代谢紊乱伴随智力低下及肝脏发育障碍）、生物素酶缺乏和枫糖尿症（一种导致身体和智力低下的先天性代谢障碍，可以分别用生物素药物剂量或限制饮食中支链氨基酸的方法治疗）（表 12-9）。这七种疾病发病率范围从 1/2 600——高发病率的镰状细胞贫血（在黑人中为 1/44）到 1/250 000——低发病率的枫糖尿症。每个新生儿筛查试验的费用为 27 美元。

表 12-9 密歇根新生儿筛查程序

疾病	遗传方式	发病率	筛查	治疗
PKU	AR	1/10 000	快速流式荧光计检测(RFA)	限制苯丙氨酸摄入
先天性低甲状腺素症	散发	1/4 000	T4 免疫组化	甲状腺素替代
镰状细胞贫血	AR	1/2 600(1/400 黑人)	高压液相色谱	严密的临床监护,青霉素预防性应用
先天性肾上腺增生(CAH)	AR	1/14 000	17-OH 黄体酮免疫荧光分析	皮质醇
半乳糖血症	AR	1/50 000~1/100 000	半乳糖和半乳糖-1-磷酸的 RFA	大豆代乳品,限制半乳糖摄入
生物素酶缺乏	AR	1/50 000~1/100 000	酶分析(比色法)	生物素药物剂量法
枫糖尿症(MSUD)	AR	1/200 000~1/250 000	抑菌实验(检测亮氨酸)	限制支链氨基酸摄入

(二)具有成年发生遗传病风险的个体症状出现前的筛查

对于成年发生的严重且可以预防的疾病如:心血管病、癌症等易感基因的鉴定,增加了遗传筛查有患病风险个体识别的可能性。一个有说服性的例子是第十一章讨论的 *BRCA1* 和 *BRCA2* 基因,这些基因的突变是多发性乳腺癌家族(含有四个或四个以上的受累近亲)中大多数腺癌的原因,同时在这些家族中也有较高比例同时发生乳腺癌和卵巢癌。虽然在这两个基因中发现了许多不同的突变,但在北欧犹太人群中是 *BRCA1* 基因发生的两个突变和 *BRCA2* 基因发生的一个突变是大多数家族性乳腺癌及卵巢癌的原因。在这个人群中三个突变等位基因联合发生的频率大约为 2.5%,使这种疾病成为在任何群体中常见的严重遗传疾病之一。正如第十一章讨论的那样,在一般化的遗传筛查计划进行之前,几个主要问题有待于解决。然而我们有理由盼望筛查 *BRCA* 的经验也将运用到类似的情况,涉及具有高发风险的严重的成年发病的其他基因突变的筛查。

(三)隐性遗传病携带者的筛查

为建立一种低成本的隐性遗传病携带者筛查项目,若干标准应是:(a)该疾病临床上症状严重,足以警惕做这样的筛查;(b)当集中努力进行筛查时,可确定高危人群;(c)一种廉价的实验,具有足够的敏感性及特异性;(d)当携带者试验检出高风险疾病时,一种确定的试验可用以进行特异诊断已鉴定的个体;(e)为有发病风险的夫妇提供有效的生育决策。携带者检测试验最成功应用的例子是在北欧犹太人群中筛查 Tay-Sachs 病的携带者。Tay-Sachs 病是一种缺乏 *N*-乙酰-β-D-氨基半乳糖苷酶(己糖苷酶 A)的 α 亚单位的常染色体隐性遗传病。这种酶的缺乏导致神经元中神经节苷脂 GM_2 积累,引起严重的神经退行性疾病,导致失明,神经元功能丧失,患者在 2~4 岁时死亡,无法治疗。尽

管在一般人群中该疾病携带者低于0.3%,北欧犹太人群中3%为Tay-Sachs病携带者;因此,筛查计划能集中在限定目标的这个人群中进行。一种廉价的试验可用来测量血中不耐热的己糖苷酶A(hexA)活性;利用测量白细胞中hexA的活性来确定携带者的情况(因为怀孕时血清中hexA的活性增高,筛查孕妇时须用白细胞检测)。北欧犹太人口中98%的Tay-Sachs病患者发生了三个突变,因此要有可用的方法进行确切的DNA诊断。产前诊断对Tay-Sachs病是有用的,以便后代中有风险的夫妇选择对怀有的受累胎儿终止妊娠。在过去的25年中,该试验在北欧犹太人口中的广泛应用实际上已在这个群体中减少了Tay-Sachs病的发生。大约有一百多万人作了筛查,Tay-Sachs病在这一人群中的发病率已降低了95%。

其他适合做这样筛查的疾病的例子有发生在地中海人群中的β地中海贫血,发生在亚洲人群中的α地中海贫血,以及发生在黑人群体中的镰状细胞贫血和血红蛋白病。在Sardinia地区及意大利部分地区所进行的β地中海贫血筛查项目,使受累妊娠的夫妇选择终止怀孕,已将此病的发病率降低了95%。

囊性纤维化(CF)是发生在白人中的最常见的常染色体隐性遗传病。在美国活婴中的发病率大约为1/2 500,携带者频率大约为1/25。目前已描述了*CFTR*基因中的600多种突变,虽然大多数突变是非常罕见的。北欧血统中大约70%的囊性纤维化染色体及南欧血统中约50%的CF是由于一个单突变ΔF508。花费大约150美元,筛查70个包括ΔF508在内的常见突变,可鉴别90%北欧人的CF携带者。具有讽刺性的是,北欧犹太人口中CF的发病率较低,仅分析5个突变就可鉴别97%的CF携带者。在一些研究中心,CF试验与Tay-Sachs病筛查项目一起进行。在黑人群体中,可检测到61%的携带者。对疑有的受累胎儿提供产前诊断。与Tay-Sachs病不同,CF是可以治疗的。在过去的几十年中,治疗方法的改进使现在患有CF的个体寿命明显延长了约30~40年。因此,CF携带者遗传筛查已经变成一个更复杂的心理学及伦理的问题。

在健康和社会政策上的意义

遗传筛查试验的扩大应用提出了许多健康和社会政策上的重要问题。遗传筛查曾经主要是医学遗传学家们在学院才涉及的领域,现在已成为普通医生和一些商业性实验室的常规实验手段。显然,在遗传学领域及其在卫生保健的应用将需要优秀的经特殊培训的卫生专业人才及相应的政策来进行这方面的工作。对成年期发生的严重疾病(如前面提到的乳腺癌)的遗传筛查的进展,将产生一些新的重要问题。未来的遗传试验将要检测主要精神疾病的易感性,如精神分裂症、狂躁-抑郁症等将会陷入更深的困境。针对降低遗传病发病率及影响的公共健康策略,将与维护个人自主权的决策需要之间产生矛盾。这些问题将在第十四章中进一步讨论。

四、遗传疾病的产前诊断

产前诊断的目的是当未来父母生育受累子女的风险极高时,给他们提供一种无受累子女的安全保证,这并不等于保证生育正常子女。产前诊断使人们在一定程度上改变某

种特定疾病的风险概率，但不是指所有可能发生的出生缺陷或遗传病。

产前诊断是应用基础科学技术解决临床问题的范例。20世纪60年代，由于产科技术、人体细胞培养、细胞遗传学及随后的分子遗传学技术的进展推动了诊断技术的发展。产前诊断运用损伤性（羊膜穿刺术、绒毛采样、经皮采血）和非损伤性方法（超声诊断技术）。

（一）羊膜穿刺术（amniocentesis）

羊膜穿刺术即从羊水中（图 12-17）抽取少量液体（通常为 20ml）。羊水是一种渗出液，发育中的胎儿在妊娠早期浸泡在其中，其成分反映胎儿细胞外液的内容，其中含有从皮肤、消化道、呼吸道及泌尿生殖道内壁上皮脱落的胎儿细胞。在妊娠晚期，羊水大部分为胎儿的尿液，其容积也随之增加。羊膜穿刺术必须尽早操作是有益的，例如可使特异诊断试验得以实施，其结果可用来决定在安全、合法的前提下是否继续或终止妊娠，晚期操作仍是安全的，例如取出少量的羊水不影响胎儿的继续发育。虽然一些研究中心试想把这种方法提早到 11～14 周，但通常还是在妊娠 16 周进行。

在进行羊膜穿刺术本身操作前，必须首先进行关于手术的适应症、好处、风险性及禁

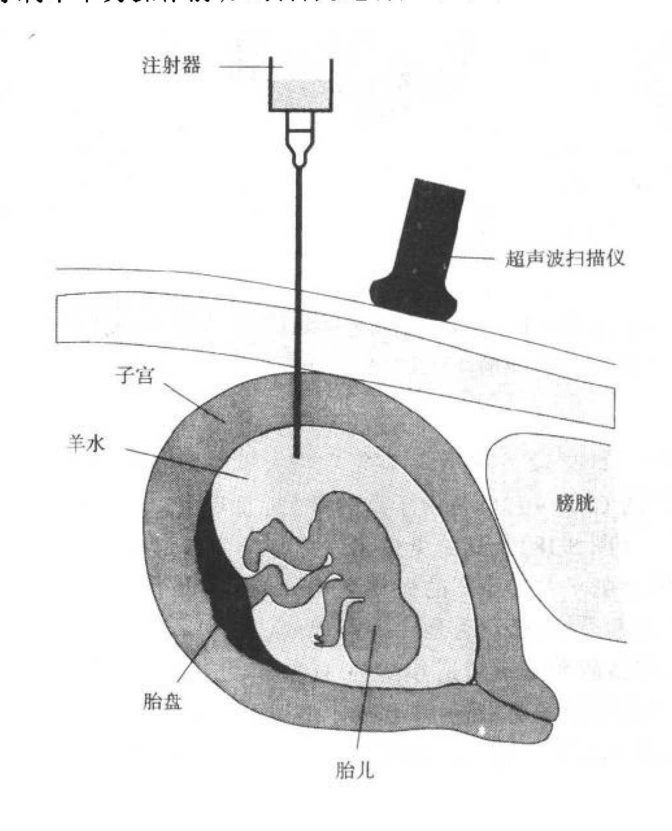


图 12-17 羊膜穿刺术示意图。

忌等准确的辅助性遗传咨询。通常要做超声图像诊断来确定胎儿在子宫中的部位(图 12-18),明确实际的妊娠时间及判断是否为多胎妊娠(双生子)。羊膜穿刺术在门诊进行,采用局部麻醉。取出的羊水可直接分析特异蛋白如甲胎蛋白,也可以培养或做风疹病毒的感染分析。更重要的是羊水中的细胞可离心收集直接分析或进行培养后再分析。液体中许多细胞是无活力的,但可用来建立羊水细胞培养物。这些培养物可运用荧光原位杂交(FISH)、常规细胞遗传学技术、生化分析或遗传分析等进行细胞学检测。

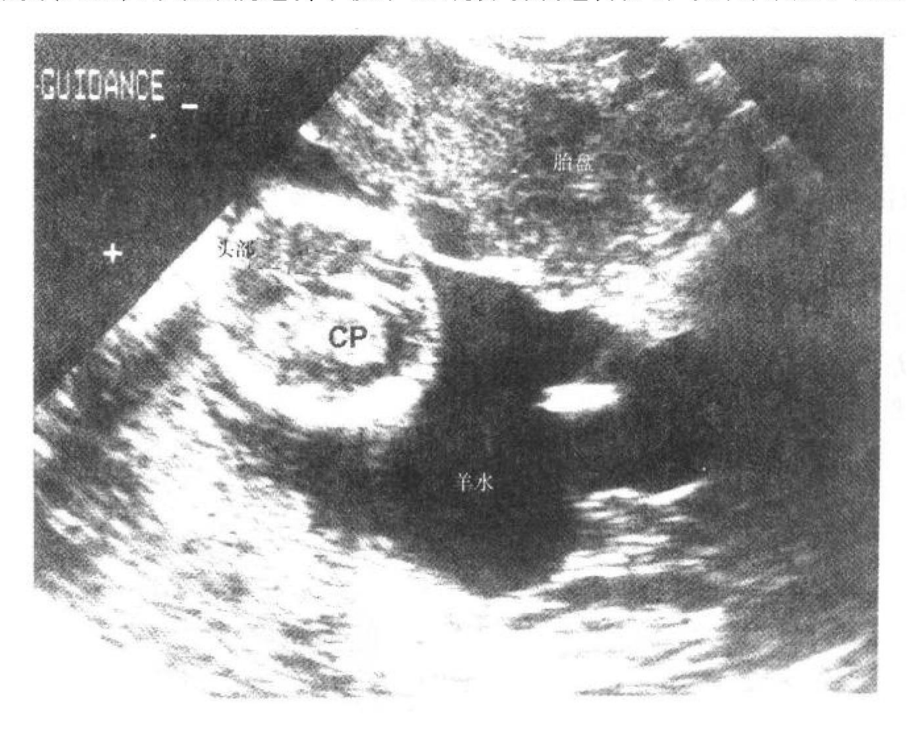


图 12-18 妊娠 16 周行羊膜穿刺术时的超声影像图。可见羊膜穿刺针从右上方进入羊膜腔;亮区是从针尖发出的反射波。前置胎盘。左侧可见胎儿头部脉络丛。

羊膜穿刺术指征如表 12-10 所示。在密歇根大学医院 90%以上羊膜穿刺术的实行是为检测染色体。大约 65%~70%是由于高龄孕妇,通常限定为 35 岁以上的妊娠妇女,这个年龄 Down 综合征(图 8-18)及其他染色体三体型的风险显著增加。在这个年龄段胎儿染色体畸变与羊膜穿刺术引起流产的概率是相同的,约为 1/200。对任何 35 岁以上孕妇行羊膜穿刺术(或绒毛采样)已认为是一种常规的监护。在许多研究中心,提供羊膜穿刺术的年龄已降低,许多研究显示最适年龄为 31 岁。大约有 20%~25%羊膜穿刺术是为有异常母体血清多重标记筛查、提示怀有三体型胎儿的妇女进行的。其他的染色体指征包括:以前出生的孩子有染色体三体型的既往史,染色体异常家族史,或已知的出现平衡易位的携带者。

表 12-10 羊膜穿刺术指征

1.高龄孕妇
2.异常的母体血清三倍体标记筛查结果
3.染色体异常家族史
4.神经管缺损/异常母体血清甲胎蛋白
5.生化检查或 DNA 分析检测的孟德尔式遗传疾病

羊膜穿刺术也可在母体血清中甲胎蛋白增高时进行,或在以往出生的孩子有神经管缺损既往史时施行(图 12-19)。由于 Down 综合征和神经管缺损相对高发,所有羊膜穿刺术常规利用细胞做染色体研究,利用羊水液体检测甲胎蛋白。羊膜穿刺术最后一个指征是有易于进行 DNA 或生化诊断的孟德尔式遗传病家族史和/或有来自一对配偶有受累子女风险的筛查试验证据。有 400 种孟德尔式遗传病适合进行这样的诊断且病种在不断增加。尽管许多疾病相当罕见,但其他诸如镰状细胞贫血、β 地中海贫血、囊性纤维化和脆性 X 染色体综合征等疾病是相当常见。

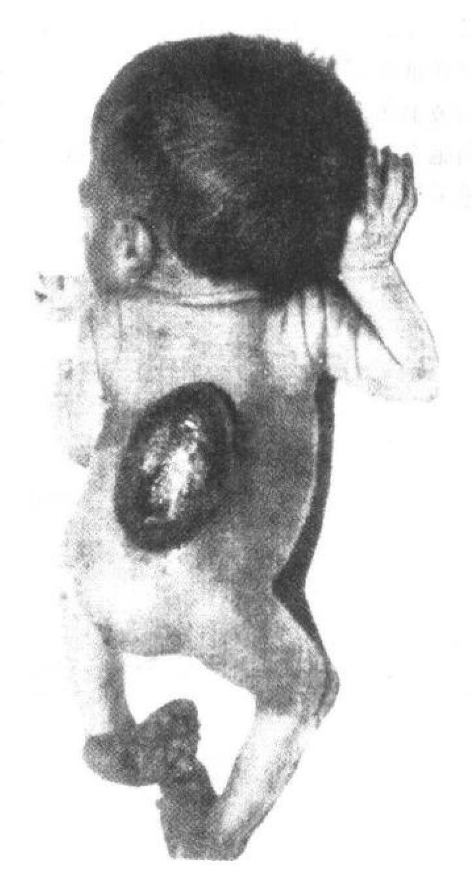


图 12-19 神经管缺损的婴儿。从胸廓下部至骶骨上部脊髓脊膜突出。

妊娠中期(16 周)行羊膜穿刺术的危险性大约有 0.5%或更少。这包括对胎儿和/或母体的感染和损伤。事实上这些危险性可能更小。重要的是要注意:几个大规模的对照研究显示:已经施行羊膜穿刺术的妇女与未施行该术的对照组妇女相比,其流产风险的增加无明显统计学意义。但早期施行羊膜穿刺术(14 周以前)却有显著增加。

在一个实验中心,没有取出羊水或没有从羊水中培养出细胞的可能性低于 0.1%。羊膜穿刺术后产生的问题包括:延误了能够从羊水中培养出细胞的时间,无细胞生长和/或细胞污染,能够混淆实验结果的母体细胞对羊膜细胞的污染,染色体嵌合或其他实验错误。幸运的是,羊膜穿刺术已证实是一种高度准确、可信的诊断技术,而其误差率在 1%以下。然而重要的是须牢记决定的关键(可能是终止妊娠)是以实验诊断为基础而非患者的体格检查。医生们试图在临床上避免出现这些情况,而产前诊断一般就是必不可少的。

(二) 绒毛采样(Chorionic Villus Sampling)

另一种产前诊断的方法是绒毛采样(CVS)。这种技术利用一个精细而有弹性的聚乙烯导管在超声探测仪监测下由子宫颈进入子宫腔,吸取少量绒毛组织,这些绒毛组织最终会发育成胎盘(图 12-20),也可穿过腹腔进行,两种方法的安全性及效率是相同的。胎儿的绒毛组织必须仔细地与母体组织分离开,并且可以用多种方法检测。胎儿的滋养层细胞是外胚层组织,是进行有丝分裂的快速分裂细胞,所以这些细胞不需进行培养就可检查核型。然而,由于嵌合的原因使核型分辨率不高,通常要从绒毛组织的间质中心培养细胞用于细胞遗传学分析。生化分析法也可用于培养细胞。DNA 分析法既可直接用绒毛细胞,也可用培养的绒毛细胞。CVS 样本中 DNA 的产量比从 20ml 羊水中获得的细胞多,常可直接进行 DNA 分析而无需先做细胞培养。密歇根大学医院 90%以上行绒毛采样的孕妇为高龄孕妇。

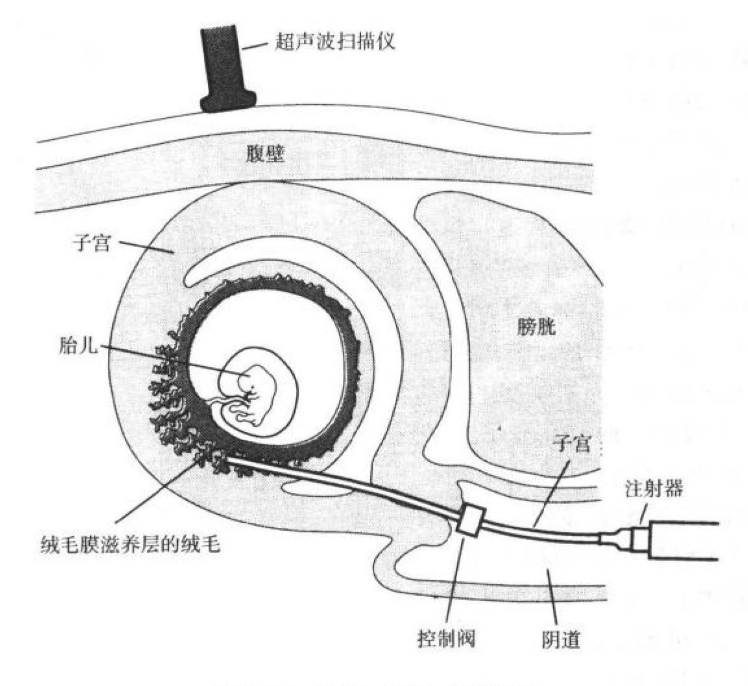


图 12-20 绒毛采样技术示意图。

绒毛采样通常在怀孕 9~12 周时进行。依据实验结果若选择终止妊娠,在时间上比妊娠中期做的羊膜穿刺术在医疗上更安全,在感情上也易接受。羊膜穿刺术和绒毛采样在诊断准确性和花费是相同的(在密歇根大学医院大约须 2 000 美元,包括咨询费、超声检查、操作、医院及实验室的各种费用)。CVS 的缺点是不能诊断神经管缺损,羊膜穿刺术也不能;行 CVS 的孕妇必须在 16 周时做 MSAFP 试验及超声检查。尽管与羊膜穿刺术相比行绒毛采样的孕妇在操作失败及胎儿流产的风险方面略有增加,但 2278 例 CVS 的预期结果表明该技术对早期产前诊断是安全而行之有效的。1996 年 6 月全世界 81 个研

究中心进行了 200 000 例绒毛采样术，胎儿流产率大约为 3%，在一些大的研究机构接近 2%，或比没有行 CVS 的孕妇流产率大约高出 1%以上。全世界大约有 4.5%行 CVS 的孕妇需做选择性流产。尽管有报道行 CVS 后胎儿会出现四肢短小缺陷，但在 9 周以后出现该缺陷的频率与群体中发生频率相同。

有趣的是 CVS 样本中发生染色体异常的频率比羊水样本高出 2 至 5 倍，大约是出生活婴的 2 倍。这与妊娠过程中对核型异常胎儿的选择性自然流产行为是一致的。虽然在绒毛采样前大多数严重染色体异常胎儿会发生自然流产，这些资料表明胎儿流产会发生在妊娠的全过程。

(三) 经皮采血(Percutaneous Blood Sampling)

胎儿的血液可在妊娠大约 18 周时采到，在超声波探测仪监视下将针插入脐带中抽取。经皮采血(PUBS)可用于检测继发于 Rh 不相容同种免疫疾病，胎儿血液异常，感染，及对超声波检测出的先天畸形做快速核型鉴定(见下文)。对培养 48~72 小时后的胎儿淋巴细胞做细胞遗传学分析，可以在利用超声图像检出的先天畸形中检测出威胁人生命的大多数染色体异常。这种方法造成胎儿流产的风险大大提高，比原有的风险高出大约 2%以上。

(四) 超声检查(Ultrasonography)

超声图像通常用于计算妊娠时间，监测胎儿生长状况，检测多胎妊娠。也可用于检查妊娠 18~20 周的男性胎儿的生殖器来确定胎儿性别。超声诊断对那些以多基因或多因子方式遗传且不适合做特异生化及 DNA 分析的解剖学缺陷是必不可少的。虽然常规超声可以诊断大多数心脏、肾脏、膀胱等畸形，但有某种明确的畸形家族史、异常 MSAFP、或有可疑异常超声波时可做靶向超声波检测，诊断许多先天畸形的总体敏感性及特异性在 95%以上(图 12-21)。胎儿的超声波心动扫描为检测常见的由于先天畸形而致死的先天性心脏病提供了重要的辅助性诊断方法。正如第八章讨论的那样，主要的染色体异常往往伴有严重畸形。反过来，超声检查出的多重先天畸形也提示存在染色体异常。当有此疑问时，通常在羊膜穿刺术后进行核型分析。如果需要进行快速细胞遗传学诊断，而且会在很大程度上影响对胎儿的医疗处理时，可以用经皮采血法。

(五) 新进展

遗传病的诊断技术已经发展到在受精后六天胚胎着床之前即可进行。这些技术使孕妇既可避免分娩患遗传病的婴儿，又不必进行流产术，在器官形成之前就对受累胎儿加以识别以便进行早期基因治疗。孕前诊断包括选择有正常等位基因的配子进行体外受精和移植。这种方法需要卵母细胞的显微操作来分离和分析第一极体，或从 4~8 个细胞的囊胚中取出一个单细胞；用这种方法可诊断囊性纤维化。这两种方法都需要体外

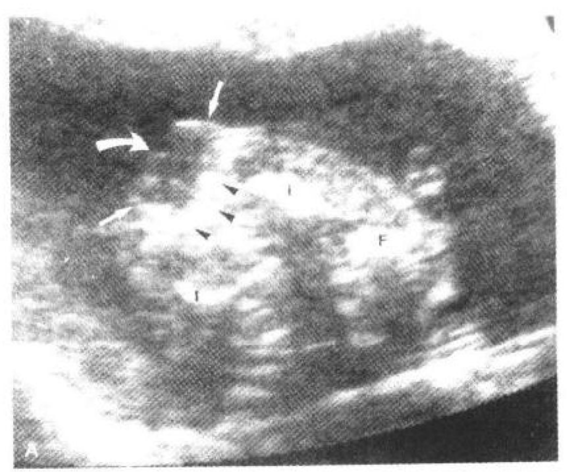






图 12-21 人胎儿的超声波检查。A:伴有神经管缺损的 17 周龄胎儿的超声诊断图像。这是穿过骶关节和髂嵴水平的胎儿腹腔横断面,用 I 表示。黑色箭头表示三个脊柱骨化中心,表示脊柱后部呈八字形向后展开。白色箭头和白色弯曲箭头表示脊髓脊膜突出的囊状团块。F 表示腿部。B:继发于后部尿道瓣膜泌尿器梗阻的 18 周龄胎儿超声诊断图像。这是穿过胎儿腹部的纵断面,T 表示喉部,位于右上方。前方和后方两极为双侧肾脏,用黑色箭头表示。可见由肾内收集系统扩张产生的拉长影像。C:伴有脐突出的 18 周龄胎儿超声图像。这是穿过胎儿的矢状断面,H 位于右上方表示头部,A 表示腹部,用白色弯曲箭头表示。

受精。第三种方法是对 100～200 个细胞的胚囊进行滋养外胚层的活检;这种方法需要自然受孕及从子宫中收集受精卵。生殖技术、显微操作技术和利用 PCR、FISH 等对单个细胞进行遗传学分析等技术的进一步改进导致上述方法的进步。但现在过高的费用及技术本身的需要限制了一些很专业的研究中心对这些技术的应用。

应用胚胎镜对早期胚胎的直接观察,一般在妊娠 5～8 周后进行。然而此技术正在实验阶段,其风险还有待确定。

羊膜穿刺术及绒毛采样是高度准确和安全的方法已成为常规产科检查手段。但它们是损伤性方法,会对母体和胎儿带来很小但已肯定的危险。而且,这些方法用于 35 岁以下孕妇的机会较少,仅在出生多个孩子或多个患有严重遗传疾病如染色体三体型的孩子时才采用。所以迫切需要发展非损伤性方法来进行产前诊断。母体血清特异蛋白的检测,如检测孕妇神经管缺损和染色体三体型的三倍体标记筛查,是筛查试验而非确切的诊断试验 。因此,几个实验室已努力分离和分析母体外周血中稀少的胎儿有核细胞,这些细胞是进行遗传诊断胎儿染色体和 DNA 的 来源。

几种类型的胎儿细胞包括滋养层细胞、淋巴细胞、粒细胞、干细胞和有核红细胞可以早在妊娠 8 周的母体血中发现,其量极少。母体血中胎儿细胞的出现频率估计为一亿分之一。在母血中对胎儿细胞做遗传学分析的关键在于发展细胞分离技术,增加胎儿细胞数量,同时发展分析技术可以检测较少量的细胞。许多工作的焦点集中在未成熟的红细胞上,它包括成红血细胞和有核红细胞,它们在胎儿血液中大量存在,而在成人血中却很稀少。分离技术依赖于对胎儿细胞表面抗原的识别。许多技术应用荧光激活细胞分类法(FACS),可以获得 10^4 ～ 10^5 倍的信号。磁力激活细胞分离系统,利用磁珠黏附于相应表面抗原的抗体上,是一种较简单而费用又低廉的分离技术。由于细胞分离技术不可能完全把胎儿细胞与母体细胞分离开,因此从分离物中确定胎儿基因序列是必不可少的。在历史上,这是利用 PCR 技术确定 Y 染色体的序列完成的。父系遗传的 DNA 多态检测已被运用于胎儿细胞。利用如下的方法,在对选定区域的 DNA 作 PCR 扩增后可以作遗传缺陷的分析。目前,即使在含有大量胎儿细胞的标本中,一般也需对其中存在的大量母体细胞排除携带突变基因的孟德尔式疾病的诊断,这是因为母体细胞的 DNA 优先被扩增且很难与胎儿细胞 DNA 区别。然而在分离了个体的胎儿细胞后,已对镰状细胞贫血及 β 地中海贫血成功地作出诊断。最初用于定向检测非整倍性的细胞遗传学诊断可采用荧光原位杂交(FISH)的方法,在细胞分裂间期核中识别染色体数目异常。现在正研究在母体循环中,胎儿细胞的细胞遗传学诊断方法的特异性及敏感性。

(六) 产前诊断中 DNA 分析的应用

许多基因的表达具有组织特异性。这就限制了对那些基因产物仅在羊膜或绒毛膜细胞表达的疾病进行生化诊断。基因表达产物的缺陷仅在特定组织如肝脏(如鸟氨酸转氨甲酰酶)表达,不能用这种生化技术作诊断。因为所有体细胞中的 DNA 实质是相同的,所以在 DNA 水平检测突变可克服这种局限。有几种技术可用于 DNA 分析(表 12-11)。

表 12-11　应用 DNA 分析进行诊断

1.已知点突变、拼接突变、小片段插入或缺失的检测
  a.限制性内切酶分析检测发生于限制性位点的突变
  b.等位基因特异寡核苷酸杂交
  c.PCR 扩增和直接自动测序
2.PCR 扩增检测未知突变
  a.单链构象多态分析
  b.异源双链核酸分析
  c.错配切割分析
  d.DNA 测序
3.大片段插入、延长、缺失或其他主要结构重排的检测
  a.DNA 印迹分析
  b. PCR 扩增
4.利用基因内 DNA 多态连锁分析诊断具有多个致病突变及家族内未知特异突变的疾病

为检测已知的点突变、拼接突变、小片段缺失或插入可以对相应的区域作 PCR 扩增并用几种技术加以分析。如果单个碱基突变引起限制性酶切位点的改变,可以很容易地用限制性酶分析法对疾病作出产前诊断。一个范例如引起镰状细胞贫血的突变导致了 *Mst* Ⅱ 位点的丧失(已在第六章讨论)。现在比较常用的技术是等位基因特异性寡核苷酸(见第五章)。这种技术尤其适用于那些由一个或多个突变引起疾病的大多数病例,如 α_1 - 抗胰蛋白酶缺乏症、Tay-Sachs 病、镰状细胞贫血和在家族中已确定为特异基因突变的疾病。在 β 地中海贫血的病例中有多个突变,这种方法对特定人群(例如:某些地中海人群中 IVSI100 的拼接突变和撒丁岛人群中的 β^{39} 无义突变)中普遍存在的小数目异常突变仍然适用。同样,对于 *BRCA1* 和 *BRCA2*,虽已确定有 100 个突变,但在北欧犹太人群中大多数仅有三个突变。最后,随着微阵芯片 DNA 自动测序技术的发展,直接测序已成为 DNA 诊断技术的常规方法,这种方法尤其适用于已知有限数目突变引起疾病的大多数病例。

未知突变可用外显子 PCR 扩增及随后的物理、生化方法来检测,已在第五章中讨论,包括单链构象多态(SSCP)分析、异源双链分析和错配切割分析法。公认的突变已通过直接测序得到证实。尽管这些方法并非普遍用于产前诊断,但对于在出生后确定早先未知的引起疾病的突变是有价值的。

大片段的缺失、延长、插入或其他主要结构重排可以用 DNA 印迹法和 PCR 方法检测。目前,多重 PCR 分析对 Duchenne 肌营养不良的产前诊断最为适用,在该疾病中有 60%以上的病例是在这个很大的基因中发生缺失的结果。DNA 印迹杂交分析也适用于由基因倒位引起的 α 地中海贫血和 A 型血友病的诊断。DNA 印迹分析可作为选择方法对 X 脆性综合征和先天性强直性肌营养不良三联体重复动态突变进行诊断,而 PCR 方法则对检测 X 染色体脆性综合征的前突变和 Huntington 病中三联体动态重复序列的检测更为有用。

利用基因内 DNA 多态进行连锁分析需要对家系作研究。这是对单基因病进行产前

诊断的最适用的一般技术，这些疾病具有多个致病基因突变和家族中未知的特异突变。已应用这种技术的疾病有：血友病、家族性高胆固醇血症、抗肌萎缩蛋白基因无缺失的Duchenne肌营养不良家系，α_1－抗胰蛋白酶缺乏症，成人多囊肾等。

最后，如果多重特异突变产生不完全构形、无功能蛋白产物，蛋白质截断鉴定法对诊断如Ⅰ型神经纤维瘤和家族性多发性结肠息肉等是有效的。运用DNA分析法（表12-12）可诊断的疾病种类正在迅速增加，包括血红蛋白病、甲和乙型血友病、α_1－抗胰蛋白酶缺乏症、苯丙酮尿症、Duchenne肌营养不良、囊性纤维化、以及一些迟发的疾病如成人多囊肾、Huntington病等。在许多病例中可将上述技术结合用来作出诊断。

表12-12 利用DNA分析法适于作产前诊断及症状前诊断的选择性单基因疾病

常染色体显性疾病
- 强直性肌营养不良
- 成人多囊肾
- Huntington病
- Ⅰ型神经纤维瘤
- 家族性乳腺癌

常染色体隐性疾病
- 镰状细胞贫血
- α地中海贫血、β地中海贫血
- 囊性纤维化
- 苯丙酮尿症
- α_1－抗胰蛋白酶缺乏症
- Tay-Sachs病

X连锁隐性遗传疾病
- 甲、乙型血友病
- Duchenne肌营养不良及Becker肌营养不良
- X染色体脆性综合征
- 鸟氨酸转氨甲酰酶缺乏症

（七）伦理学问题

有关临床遗传学的伦理问题将在第十四章进一步讨论。由产前诊断产生的主要伦理问题与流产有关。如果一对夫妇冒很小但确定的危险作产前诊断，他们对受累胎儿有终止妊娠的义务吗？必须注意的是约少于5%作产前诊断的胎儿患有正在咨询的疾病。因此对绝大多数作产前诊断的妇女来说，其结果是让她们安心没有妊娠疑有疾病的胎儿。而且，一个妊娠受累胎儿的妇女应有机会考虑其实际情况作出最佳决策而非先前提到的义务。

第二个问题是如何处理不需要或意外的信息。因为无论羊膜穿刺术的最初指征是什么，所有的标本都要作核型和甲胎蛋白分析，所以可能得到意外的信息。然而考虑到许多其他医学诊断实验也会发生此问题，许多遗传咨询家们感到应当提供全面的、适当

的、支持性遗传咨询。如何处理双胎妊娠中,特别是双生子患有严重遗传疾病表现不一致的情况? 例如,其中一个胎儿患有 Down 综合征而另一个则没有? 有一些这样病例的报道,选择性流产受累胎儿,而让正常胎儿出生是可能的。这些情况再次强调了产前诊断的局限性:除非是在罕见情况下,终止一个患有遗传病的胎儿生命是可能的,而治疗患病胎儿不可能。对于妊娠患有严重或致命的遗传病胎儿的夫妇来说要面对痛苦,这种选择比任何其他选择更人道,其结果也令人满意。

(八) 产前诊断对基因频率和发病率的影响

产前诊断及对患病胎儿作的流产术对其他致死性遗传病的基因频率无影响。没有产前诊断患病胎儿也会在出生前死亡。致死性常染色体隐性遗传病对基因频率产生的影响,取决于有风险的携带者是否有很多(很少)的孩子(如果可以避免生育患病子女),这样把较多(或较少)的变异等位基因传递给后代。相反,产前诊断和选择性流产对患有非致死性常染色体显性遗传疾病(Huntington 病和家族性乳腺癌)的胎儿的基因频率会有巨大影响。

产前诊断对发病率的影响取决于是否对所有具有风险的孕妇或仅对已有患病孩子的孕妇提供产前诊断。例如:神经管缺损在一般群体中的发病率约为千分之一,如果一个妇女已生育了一个患此病的子女,其再发风险大约为 2% ~ 5%。产前诊断及对受累胎儿做选择性流产如果仅针对已生育一个受累子女的妇女,那么它对人群中神经管缺损总发病率的影响将极小(但这并不意味着对一些特殊家族没有情感上和医疗上的影响)。然而,如果对有风险的大多数妇女同时进行产前诊断及遗传筛查项目(例如母体血清中甲胎蛋白的检测筛查妊娠神经管缺损胎儿的高危孕妇),将对这种出生缺陷的发病率产生巨大的影响。同样,对 35 岁以上高龄孕妇行羊膜穿刺术和绒毛采样对 Down 综合征的发病率无明显影响。但是,对所有孕妇进行三倍体标记筛查及从母体血液中分离和分析胎儿细胞技术的发展将对此病的发生产生很大影响。

总之,分子遗传学、基因定位与克隆及对致病突变的鉴定技术的快速发展大大增加了产前诊断的应用范围。敏感的和辅助的遗传咨询与产前诊断相结合使患有严重遗传病孩子风险的夫妇在生育和防止中进行选择,如果他们选择了生育,会出生一个患病的孩子。

五、遗传疾病的治疗

(一) 原则

遗传病的治疗通常涉及环境操作而非遗传操作(基因治疗将在第十三章讨论)。应当牢记的是遗传疾病和医生们面对的“非遗传”疾病一样是可以治疗的。为避免遗传病的发生也须将遗传筛查、咨询、产前诊断、早期干预与治疗紧密结合。医学遗传学提供一个典范的例子将其重点放在疾病预防上而不是对已经发生的疾病进行简单地治疗。最

后，在环境水平上治疗遗传病一般包括三“R”的结合，即：限制疗法（restriction）、替代疗法（replacement）及切除疗法（removal）。

（二）限制潜在的有毒的环境因素

饮食疗法是治疗某些遗传性代谢疾病的重要组成部分。限制苯丙酮尿症病人饮食中的苯丙氨酸摄入可以预防日益加重的智力低下。在家族性高胆固醇血症患者饮食中，限制胆固醇和饱和脂肪的摄入对降低血清中的胆固醇有辅助作用。在一些其他疾病中限制某些药物或有毒物质是至关重要的。例如，葡萄糖－6－磷酸脱氢酶缺乏症患者必须避免服用氧化应激类药物如抗疟疾药物和一些抗生素。α_1－抗胰蛋白酶缺乏症患者吸烟（对每一个人来说都是非常有害的）是特别危险的，可导致肺组织破坏并引起严重的肺气肿。

（三）替代疗法

对一些遗传病替代缺乏的物质或器官是可以治疗甚至是治愈的。在甲型血友病患者中，替代Ⅷ因子对绝大多数病人是成功的。在 α_1－抗胰蛋白酶缺乏症的病例中，静脉注射人 α_1－抗胰蛋白酶可使肺泡液中含量达到足以抵消弹性蛋白酶的作用（图 12-22）。而替代缺乏的或异常的 α_1－抗胰蛋白酶并不能如人们所预想的那样，可防止某些患者继发于异常蛋白质在肝脏累积造成的肝损伤。

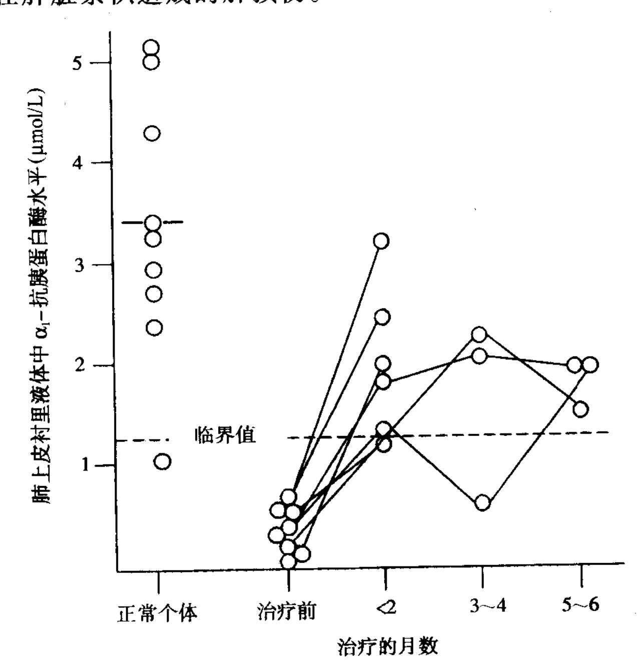


图 12-22　静脉注射人 α_1－抗胰蛋白酶后肺泡液中含此抑制剂的结果。α_1－抗胰蛋白酶缺乏症患者接受数周纯化抑制剂的注射时间如图所示。虚线代表 α_1－抗胰蛋白酶在肺上皮衬里液体中的理论估计值。

对家族性纯合性高胆固醇血症患者进行肝移植手术已经成功地纠正了代谢缺陷。与 HLA 配型的骨髓移植治愈了某些继发于腺嘌呤脱氨基酶缺乏症的免疫缺陷，已成功地用于治疗镰状细胞贫血。胱氨酸病患者的肾移植（由细胞内胱氨酸堆积引起肾脏的损伤）可纠正由此病引起的最严重的问题。

（四）切除疗法

切除有毒及受损器官的方法也已被成功应用。Wilson 病（肝豆状核变性）是一种继发

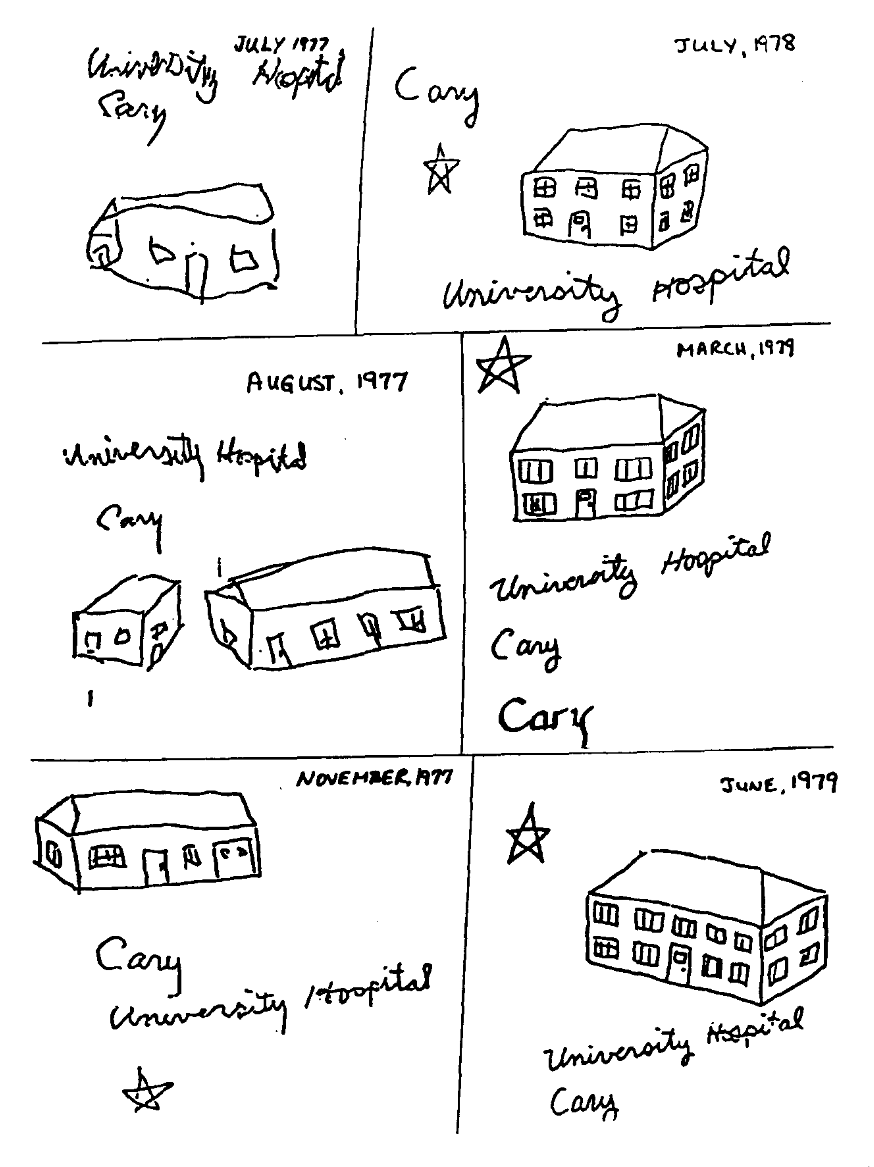


图 12-23 一位 21 岁 Wilson 病男患用青霉胺螯合法去除铜后的书法和绘画。前一部分是治疗前的画图，后部分为治疗二年后的画图。这是一种评价震颤及肌张力障碍程度的简便方法。

于铜累积所引起的肝脏和神经损伤的常染色体隐性遗传病。用青霉胺螯合铜尽管有毒性,但用于预防该病引起的损伤疗效很高,甚至在一少部分患者中纠正了肝硬化。如果早期应用该方法,还可以纠正神经损伤症状(图 12-23)。同样,在血色素沉着患者中,利用静脉切开术除去部分铁离子可预防对肝脏、心脏、胰腺等其他器官的进行性损伤。家族性多发性结肠息肉有 100%多发息肉的恶变风险,导致结肠癌的发生。切除结肠尽管不是常规疗法,但却可预防结肠癌的发生并挽救患者的生命。

除"三 R"疗法外,一些代谢疗法可用于治疗遗传代谢疾病。在继发于 LDL 受体缺乏的家族性高胆固醇血症患者中,内源性胆固醇的生物合成可被抑制剂所阻断,竞争性抑制剂 HMG-CoA 还原酶导致细胞中 LDL 受体增加和血清中胆固醇水平增高(见图 7-15)。

总之,医学遗传学特别是分子遗传学的迅猛发展实际上已使遗传学者的任务发生了从"推测者"到"修理师"的转变。产前诊断成为引人注目的焦点。临床遗传学的发展得益于基因定位、DNA 重组技术及非遗传性疾病诊断和治疗的同步发展。遗传病的诊断、治疗和预防已成为医学实践的前沿和主流。

推 荐 读 物

遗传咨询

Baker DB, Schuette J, Uhlmann W, eds Aguide to genetic counseling. New York: John Wiley & Sons, 1998.

Emery AEH, Pullen IM, eds. Psychological aspects of genetic counselling. New York: Academic Press, 1984.

Gelehrter TD. The family history and genetic counseling: tools for preventing and managing inherited disorders. Postgrad Med 1983;3:119 ~ 126.

Jones KL. Smith's recogonizable patterns of human malformation. Sth ed. Philadephia: WB Saunders, 1997.

Mckusick VA. Mendelian inheritance in man: a catalog of human genes and genetic disorders. 12th ed. Baltimore: Johns Hopkins University Press, 1997.

Murphy EA, Chase GA. Principles of genetic counseling. Chicago: Year Book Medical Publishers, 1975.

Spranger J, Benirschke K, Hall JG, et al. Errors of morphogenesis: concepts and terms. Recommendations of an international working group. J Pediatr 1982;100:160 ~ 165.

Walker AP, Genetic counceling. In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics. 3rd ed. New York: Churchill Livingstone, 1996:595 ~ 618.

遗传筛查

Andrews LB, Fullarton JE, Holtzman NA, Motulsky AG, eds. Assesing genetic risks. Implications for health and social policy. Washington. DC: National Academy Press, 1994.

Collins FC, BRCA1-Lots of mutations. lots of dilemmas. N Engl J Med 1996;334:186 ~ 188.

Frbe RW, Levy HL. Neonatal screening. In: Rimion DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics, 3rd ed. New York: Churchill Livingstone, 1996: 581 ~ 593.

Kaback M, Lim-Steele J, Dabholkar D, et al. Tay-Sachs disease-carrier screening, prenatal diagnosis, and the molecular era. JAMA 1993; 270:2307 ~ 2315

McGinniss MJ, Kaback MM. Carrier screening. In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and pratice of medical genetics. 3rd ed. New York: Churchill Livingstone, 1996:535 ~ 543.

Wald NJ, Kennard A. Prenatal screening for neural tube defects and Down syndrome. In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics, 3rd ed. New York:

Churchill Livingstone, 1996:545 ~ 562.

产前诊断

Bianchi DW. Prenatal diagnosis by analysis of fetal cells in maternal blood. J Pediatr 1995;127:847 ~ 856.

Brambati B, Chorionic villus sampling. Curr Probl Obstet Gynecol 1995;7:109 ~ 116.

Ctton RGH. Slowly but surely towards better scanning for mutations. Trends Genet 1997; 13:43 ~ 46.

D'Alton ME, Craigo S, Bianchi DW. Prenatal diagnosis. Curr Probl Obstet Gynecol Fertil. 1994; 17:43 ~ 78.

Elias S, Simpson JL. Prenatal diagnosis. In: Rimoin DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics. 3rd ed. New York: Churchill Livingstone, 1996:563 ~ 580.

Gomez KJ, Copel JA. Ultrasound screening for fetal structure anomalies. Curr Opin Obstet Gynecol 1993;5:204 ~ 210.

Shulan LP, Elias S. Amniocentesis and chorionic villus sampling. West J Med 1993;159:260 ~ 268.

Williamson B. Towards non-invasive prenatal diagnosis. Nat Genet 1996;14:239 ~ 240.

遗传病的治疗

Rimion DL, Connor JM, Pyeritz RE, eds. Emery and Rimoin's principles and practice of medical genetics. 3rd ed. New York: Churchill Livingstone, 1996.

Wealtherall DJ. The new genetics and clinic pratice. 3rd ed. Oxford. England: Oxford University Press, 1990.

第十三章　基因治疗

重组 DNA 方面的研究令人心烦意乱，并非由于现在人们所争论的危险性，而是因为所面对的事实让人们从根本上感到困惑，即控制星球生命的遗传手段能够如此轻易地被愚弄。

——Lewis Thomas，1997

本书的前几章已记述了在过去二十年中遗传学基础研究的进步及其飞速的发展。正如我们已经学习过的，许多孟德尔式遗传病的特异致病基因已经克隆，同时对患者的致病基因突变的鉴定已成为一种常规方法。

现代的临床遗传学家主要从事建立遗传病的精确诊断方法。这通常包括结合 DNA 检测来鉴定特异性突变，对患者本人及其家族成员就疾病诊断的重要性给予咨询并讨论生殖的选择。然而，分子遗传学实验技术对许多人类疾病的治疗也具有潜在的深刻影响。人们对发展中的新领域——**基因治疗**（gene therapy）寄予厚望。“基因治疗”一词一般用于描述将遗传物质（DNA 或 RNA）转移到患者体内以治疗疾病。若干单基因病已成为基因治疗的早期目标。然而这种方法最终会广泛应用于常见多因素疾病的治疗，如肿瘤和心脏病。虽然基因治疗将来会有美好的前景，但重要的是应该牢记证明这个方法成功的病例还很少，而且它目前仍处于实验医学的范围。

一、体细胞或生殖系基因治疗

图 13-1 说明基于受累组织的类型，在基因治疗的干预中有着重要的区别。在**生殖系基因治疗**（germline gene therapy）中，经修饰的治疗基因导入全身所有细胞中，或包括生殖细胞（即产生卵子或精子的卵巢或睾丸细胞）在内的另一细胞亚群，其结果是被修饰的基因能够传递给后代。而在**体细胞基因治疗**（somatic gene therapy）中，遗传修饰仅限于体细胞，对生殖系无影响。故后一种治疗仅能直接影响患者，不会将遗传操作的改变传递给后代。

常规用于实验小鼠的转基因及胚胎干细胞技术（见第五章）一般均能导致生殖系基因改变。图 13-2 所示例子中，在导入编码正常抗肌萎缩蛋白的基因后，肌萎缩模型小鼠的肌肉缺陷得到完全矫正。尽管生殖系基因治疗广泛用于实验动物，但用于人类还不能被大多数医生和科学家所接受。许多专家担心对生殖系的无意间 DNA 修饰会产生意外新的人类疾病或以不可预料的方式干扰人类进化。这种人类基因组的永久修饰所引发的道德、伦理和科学领域的难题，人们无疑将会在若干年内为之争论。虽然基因治疗目

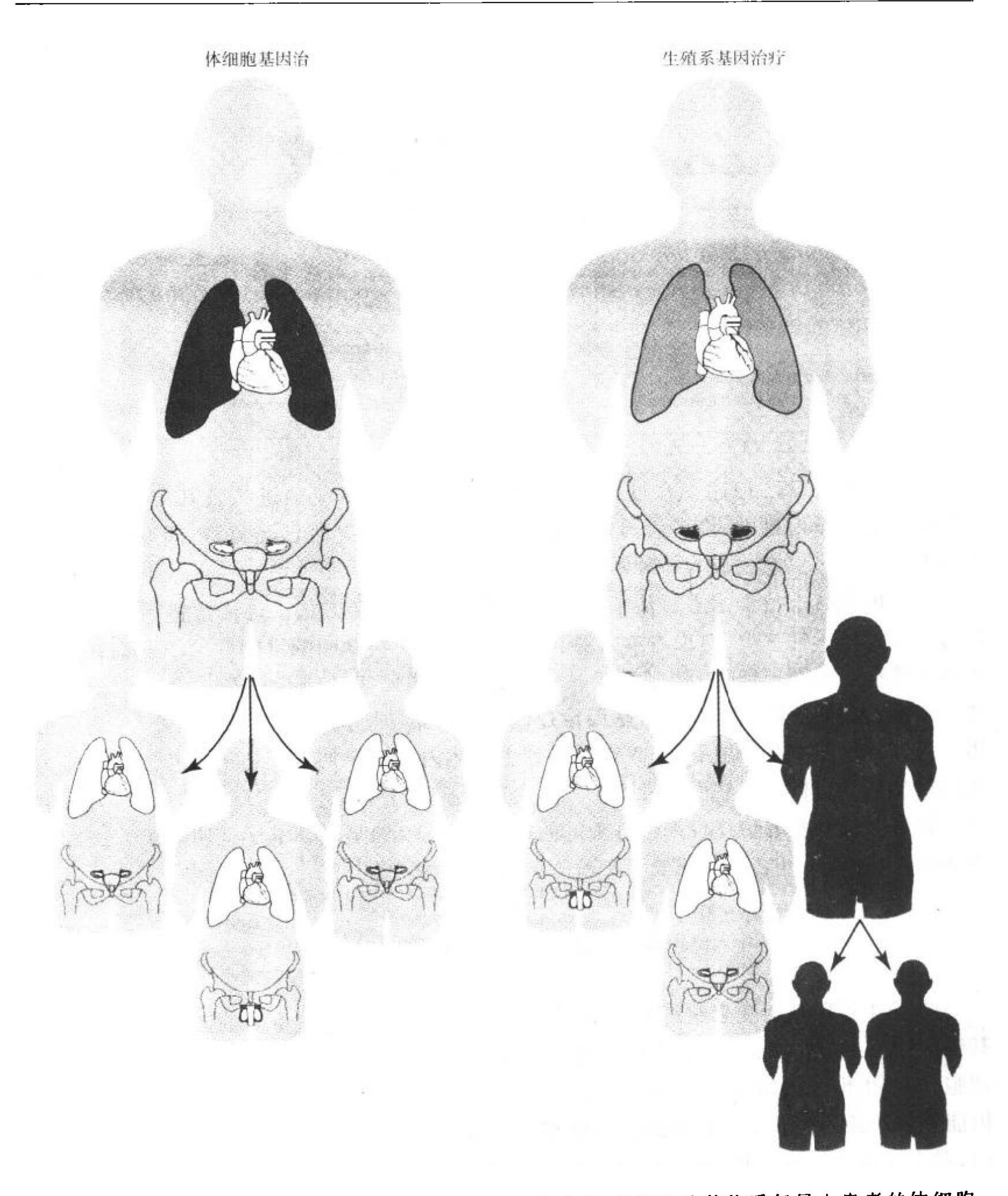


图 13-1 体细胞和生殖系基因治疗。在体细胞基因治疗中，外源性遗传物质仅导入患者的体细胞中，而不会导入卵巢或睾丸内的任何生殖细胞。因此，这种治疗对其后代无影响。相反，生殖系基因治疗是遗传性修饰生殖腺中生殖细胞，经常也修饰其他类型体细胞，故其遗传改变能够传递给患者的后代。

前尚不可行，但或许将来的科学进展能以一种可以接受的方式精确、高效地修正疾病的生殖系突变，从而解除目前对生殖系基因治疗的非官方暂禁。

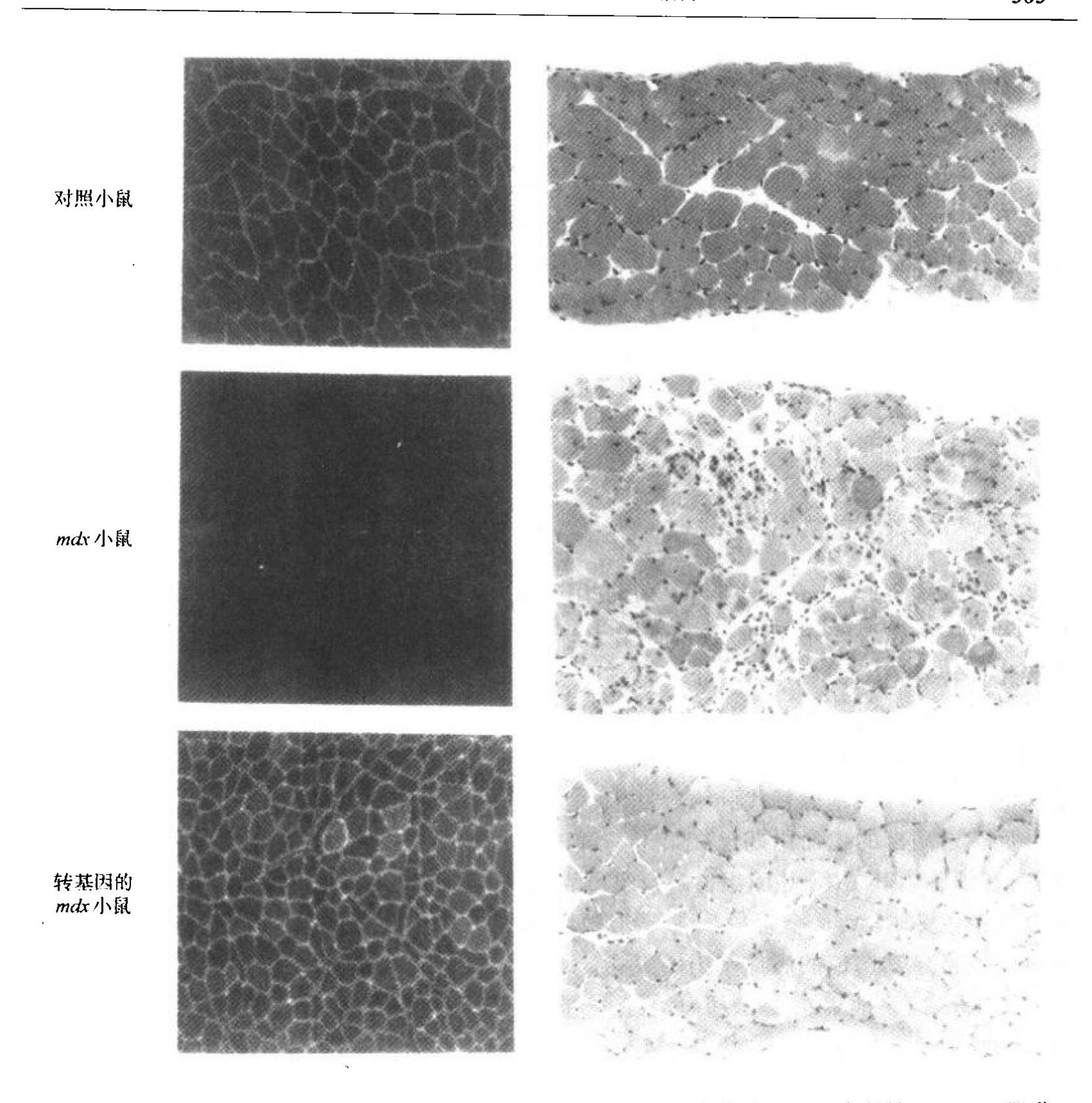


图 13-2 应用抗肌萎缩蛋白基因转基因技术矫正 *mdx* 小鼠的肌肉缺陷。*mdx* 小鼠是 Duchenne 肌营养不良(见第九章)的小鼠模型，如同 DMD 患者，*mdx* 小鼠也携带 X 染色体上的抗肌萎缩蛋白基因的突变。这类小鼠形成了许多人类 DMD 的疾病特点，包括进行性肌纤维变性及纤维化。如实验所示，在肌肉特异性启动子和增强子(来源于肌肉肌酸激酶基因)控制下，由全长鼠源抗肌萎缩蛋白基因 cDNA 产生了一个转基因小鼠系(见第五章)。然后，转基因小鼠与 *mdx* 小鼠交配，顶排为正常对照小鼠的肌肉切片，中间排来自 *mdx* 小鼠，下排来自携带抗肌萎缩蛋白基因转基因的 *mdx* 小鼠。每排左侧切片为小鼠骨骼肌抗肌萎缩蛋白的免疫组化染色，类似于人类肌肉(见第九章，图 9-23)。抗肌萎缩蛋白的染色可见于正常肌纤维外周，而在 *mdx* 小鼠中消失，但转基因后的 *mdx* 小鼠中其表达增强。右侧的显微照片为膈肌组织切片的组织结构，*mdx* 小鼠的肌肉明显变性、纤维化，但抗肌萎缩蛋白基因转基因后得到完全矫正(引自 Cox GA, Cole NM, Matsumura K, et al. Overexpression of dystrophin in transgenic *mdx* mice eliminates dystrophic symptoms without toxicity. Nature 1993;364:725 ~ 729.)。

二、外源 DNA 导入体细胞的方法

基因治疗一般需要使导入的外源基因在宿主的靶组织内表达。已应用一些方法将外源 DNA 导入细胞。用于将治疗的基因运送入其靶细胞的试剂被称为载体。载体可为病毒源性或非病毒源性。

在自然状态下的 RNA 或 DNA 病毒生活周期中外源性核酸进入宿主细胞是常见的事。这些感染体在进化中经过选择,能够有效地将其基因组导入靶细胞。也能逃避正常的宿主防御。正如我们将看到的,基因治疗学家已经开发出由几种这类病毒发展而来的许多方法。

(一) 反转录病毒

反转录病毒(retroviruses)是第一个被用作基因治疗载体的病毒系统。由于它在肿瘤基因发现中的重要作用,我们最初在第十一章讨论了这类病毒。正如你会回想到的,急性转化反转录病毒在“捕获”一个正常细胞基因后,能以一种异常的方式使该基因直接表达,导致肿瘤的发生。相对于急性转化反转录病毒破坏性的一面(Mr. Hyde),反转录病毒基因治疗载体是其有益的一面(Dr. Jekyll),类似地利用病毒结构将基因有效地导入细胞,不同的是以治疗性基因替代肿瘤癌基因。

反转录病毒的生命周期已在第十一章讨论过(图 11-3A)。它能有效地吸附并进入靶细胞,在细胞内其 RNA 基因组通过反转录酶复制成 DNA。这种前病毒 DNA 随后稳定地整合到染色体 DNA,这个阶段通常要求靶细胞能进行细胞分裂及 DNA 复制。整合的前病毒则能指导其携带的病毒基因的表达。

图 13-3 简要地描绘了在基因治疗中应用反转录病毒的一般方案,原有的病毒基因(*gag*、*pol* 和 *env*)被去除并代之以治疗性基因。因为这种人工载体序列中不再含有产生此感染性病毒所需的病毒基因,所以采用一个小策略产生重组病毒颗粒。基因治疗载体 DNA 转染一种特殊设计的生产者或**包装细胞系**(packaging cell line),这些细胞经设计携带缺陷反转录病毒表达所需的序列——*gag*、*pol* 和 *env* 等病毒基因。但这种“辅助病毒”缺少一段重要的 DNA 节段,即包装信号(Ψ),它是识别病毒 RNA 基因组并将其包装为成熟病毒颗粒所必需的。尽管包装信号在缺陷型辅助病毒内缺失,但它存在于经改造的病毒载体内。因此由辅助病毒产生的外壳蛋白不能包装辅助病毒 RNA,却能有效地包装基因治疗载体的 RNA 基因组,形成感染性的反转录病毒颗粒,并从细胞中释放出来。目前应用此种方法产生的重组反转录病毒正被收集并应用于感染靶细胞。这些颗粒带有吸附并进入靶细胞所需的所有病毒成分,在靶细胞内载体 RNA 基因组脱掉外壳,复制成 DNA 并稳定地整合到基因组。现在插入的载体 DNA 能够指导所需的治疗基因的表达。因为靶细胞不含有辅助病毒序列,所以反转录病毒感染过程在此终止,不会进一步产生感染颗粒。

反转录病毒载体具有一些独特的优点,如表 13-1 所示。或许最重要的是,它是目前

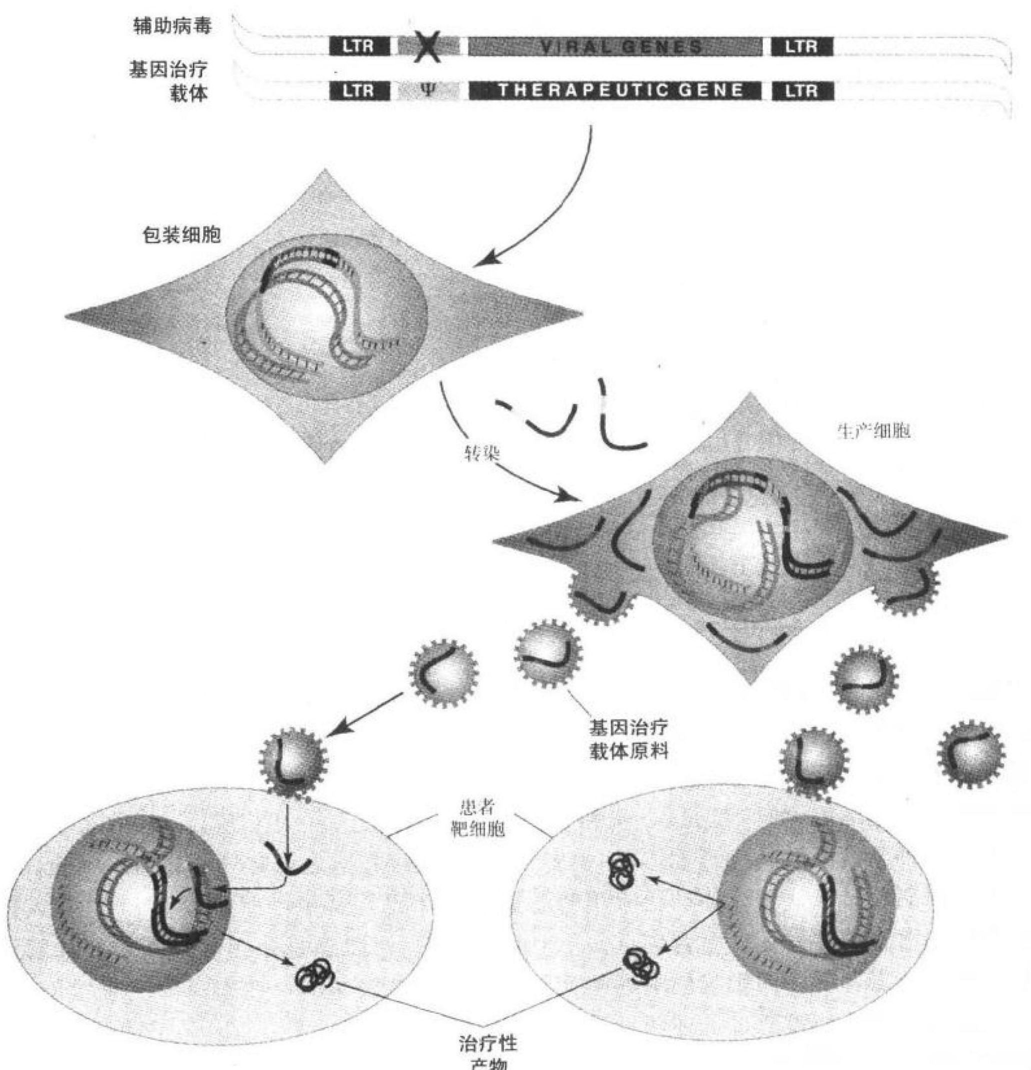


图 13-3 反转录病毒基因治疗载体的生产。基因治疗载体 DNA 包括反转录病毒两端的长末端重复序列(LTR),LTR 是高效整合入染色体 DNA 所必需的,Ψ 信号是将病毒 RNA 包装成感染性病毒颗粒必需的一段序列。这段序列在辅助病毒中被破坏,而在基因治疗载体中完整,载体内其余的反转录病毒基因序列已被所需的治疗性基因取代,基因治疗载体是应用传统的重组 DNA 克隆技术制备的(见第五章)。包装细胞是一个携带整合的辅助病毒 DNA 拷贝的稳定组织培养细胞系,这些细胞转染基因治疗载体后,可产生来自辅助病毒和基因治疗载体序列的病毒基因组 RNA 分子。但是仅基因治疗载体 RNA 含有完整的 Ψ 信号使其能够包装并以感染性病毒颗粒形式释放。当患者的靶细胞暴露于基因治疗载体时,载体 RNA 高效地导入细胞内,反转录成 DNA,并稳定地整合入染色体 DNA 中,如同自然状态下的反转录病毒生活周期(见图 11-3)。但是,因为必需的病毒蛋白基因缺失,就不会继续产生感染性病毒颗粒,基因治疗载体 DNA 只能指导治疗基因在靶细胞内表达。

惟一能够可靠地将重组 DNA 序列稳定地整合入宿主基因组的可应用系统。由反转录病毒成功转导的细胞将永久地携带治疗基因,并在细胞分裂后将其继续传递给所有的子细胞。此外,反转录病毒感染是一个非常有效的过程。在适当的条件下几乎100%地转导

表 13-1 基因治疗载体

载体	优点	缺点
反转录病毒	高效转导适当的靶细胞 长期表达整合入染色体 DNA	有插入突变产生的可能性 要求宿主是分裂的细胞 插入 DNA 大小有限
腺病毒	高效转导 靶细胞范围广泛 不需要分裂的细胞 插入突变的风险低	短暂表达 具有免疫原性 有病毒的直接细胞病理效应
腺病毒相关病毒	宿主不需为可分裂细胞 特异位点整合	若无特异位点整合,有插入 突变发生的可能 插入 DNA 大小有限
非病毒载体	无感染风险 完全合成 插入片段大小无限制	效率低 靶细胞范围有限 短暂表达

靶细胞群体。可是,这种方法也有严重的潜在缺陷。反转录病毒基因组的大小有限,因此这些载体仅能容纳中小型的基因(≤8kb)。如前所述,反转录病毒 DNA 整合到基因组一般都需要正在分裂的细胞,这种需要严重地限制了可有效转导的靶细胞类型的范围,排除了终末分化细胞,如神经元、心肌和骨骼肌细胞。最近,在 HIV 病毒(lentivirus)基础上开发了一种新型反转录病毒载体,似乎能克服这种限制,使之可转导非分裂的细胞。

或许限制反转录病毒载体广泛应用的最重要的障碍就是担心前病毒 DNA 整合入染色体会引起潜在的并发症。整合位点实质上是随机的,所以总是存在着某个重要基因被破坏或由于病毒调节序列影响使附近某个基因出现异常表达的一些小风险。事实上,人们已经证实反转录病毒通过上述两种机制产生肿瘤(见第十一章)。但是,正如我们随后将要讨论的,反转录病毒已经被用于一些人类基因治疗试验,而反转录病毒诱导肿瘤发生的风险似乎是低的。

(二) 腺病毒

腺病毒(adenoviruses)是一种人类天然的病毒性病原体,作为一种基因治疗的可能工具,人们已对它进行了深入研究。腺病毒携带双链 DNA 基因组,能够感染许多类型的细胞,仅产生轻微的人类疾病(一种普通类型是感冒)。腺病毒基因组约为 36kb,因此这种病毒作为基因治疗载体能够容纳相当大的基因。

产生重组腺病毒载体的策略如图 13-4 所示。与产生反转录病毒载体方法相似,从野生型腺病毒中去除基本的基因,既为治疗基因让出空间,又使腺病毒功能缺陷,以致其不能繁殖并在宿主体内及向治疗个体的体外扩散。如同反转录病毒,已经制备了腺病毒

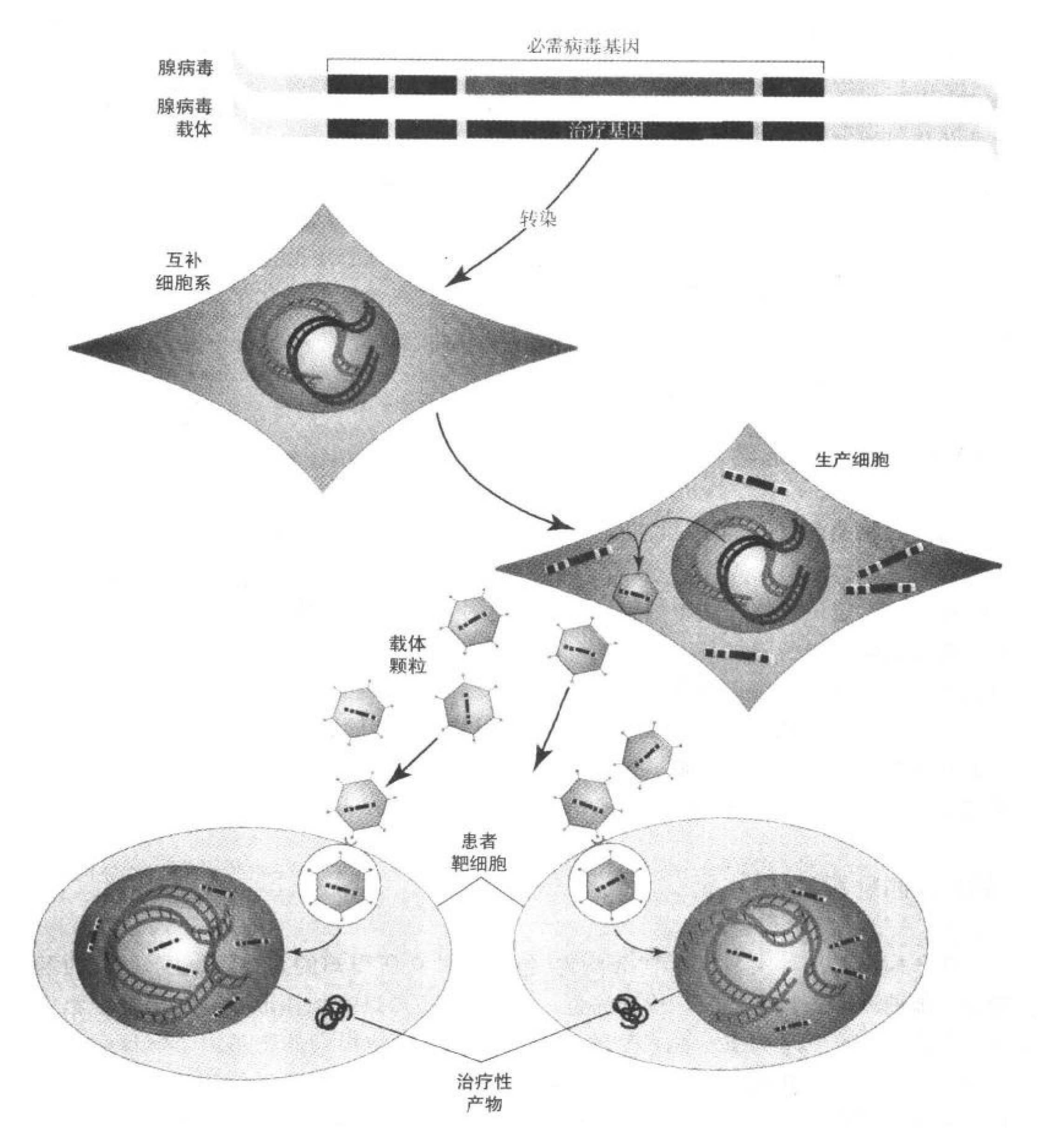


图 13-4 腺病毒基因治疗载体的产生。其总体策略与图 13.3 中描述的反转录病毒载体相似，部分重要病毒基因从腺病毒基因组中去除，为治疗性 DNA 让出空间，产生不能自我复制的重组病毒。互补的细胞系则能表达载体 DNA 中缺乏的重要病毒基因，使之能够生产包含重组病毒载体基因组的感染性病毒颗粒。载体颗粒感染的宿主细胞范围广泛，能有效地将其 DNA 作为附加体导入核内，治疗性基因序列则表达产生所需的蛋白产物。因为宿主靶细胞内不含缺失的重要病毒基因，所以不会再产生另外的病毒颗粒。

载体的特殊包装细胞系，它能够补偿重组腺病毒缺少的基本要素，使其包装成感染颗粒，并产生高滴度的病毒原液。

这个系统的主要优点是腺病毒的广泛宿主范围和感染的高效。另外，腺病毒感染不

需要分裂的细胞。病毒 DNA 以附加体形式存在,不能稳定地整合到靶细胞的染色体,后一特点既是优点又是缺点。在反转录病毒中讨论过,虽然可能因此消除了整合至某一关键基因引起的致癌风险,但附加体随时间推移逐渐丢失通常会导致短暂的基因表达,特别是在分裂的细胞中更是如此。

宿主对腺病毒的免疫反应也是此载体的一个主要限制因素。因为腺病毒是人类天然病原体,大多数个体对腺病毒的外壳蛋白已有一定程度的免疫力,或经治疗后迅速产生免疫。这种炎症反应在腺病毒载体的短暂表达中是一个常见主要因素。腺病毒引起的炎症反应本身就可能成为一种严重的并发症,正如在囊性纤维化早期基因治疗试验中至少在一个患者身上看到(见后)。对腺病毒载体进行进一步的修饰以减少其免疫原性可能会提供解决这些问题的方法。

(三) 腺病毒相关病毒

腺病毒相关病毒(adeno-associated virus AAV)是一种含有 DNA 的细小病毒,其复制需要与腺病毒共感染。AAV 复制为双链 DNA,但包装成单链。野生型病毒通常在人类 19 号染色体的特定位置整合到靶细胞基因组。这种特点使其能够稳定地整合到染色体,而无反转录病毒那样的随机基因失活的危险。不幸的是,到目前为止所检测的重组 AAV 载体似乎已失去特异性位点整合的能力。

几乎全部的 AAV 基因组都能被治疗性基因替代而去除,但 4.7kb 大小的野生型病毒携带外源基因大小仍非常有限。最近的实验表明在使用 AAV 的动物模型系统中表达水平前景乐观,但仍存在一些重要问题有待克服。

(四) 其他病毒载体

疱疹病毒是一种大型双链 DNA 病毒,包括几种众所周知的人类病原体,已作为基因治疗可能的载体进行了研究。像腺病毒一样,疱疹病毒以附加体形式存在于靶细胞中且不能稳定地整合到基因组。其大型的基因组有能力容纳很大的靶基因。疱疹病毒已被试验应用于有限的几种情况,特别是应用于将基因导入中枢神经系统。包括牛痘和流感病毒在内的其他一些病毒也作为基因治疗可能的工具正在研究,但是这些研究仍处于相当早期的阶段。

(五) 非病毒载体

由于我们已讨论的所有病毒载体都存在自身固有的问题,迄今已经开发了若干只用 DNA 而不需病毒感染辅助的方法。虽然,简单注射裸露的、纯化的 DNA 可被肌细胞高效摄取并表达其编码基因,但在其他组织中该方法效果不佳。然而,当以肌肉为靶目标或对细胞类型要求不严格时,这种方法十分有效。

目前已发展了一些增强非病毒 DNA 策略,包括将 DNA 包被脂质层的**脂质体**(lipo-

some)。在脂质体包裹的DNA上包被有针对某一类型靶细胞上的特异受体的配体，也可能提供一种进行选择性的DNA运送的有效方法。上述及其他方法能够有效地介导DNA进入靶细胞而不会引起与病毒成分有关的问题。脂质体介导的基因转移已用于一些人类基因治疗试验。

三、基因治疗的策略

如图13-5所示，将所需的基因治疗载体输送到患者体内一般有两种方法：转体内基

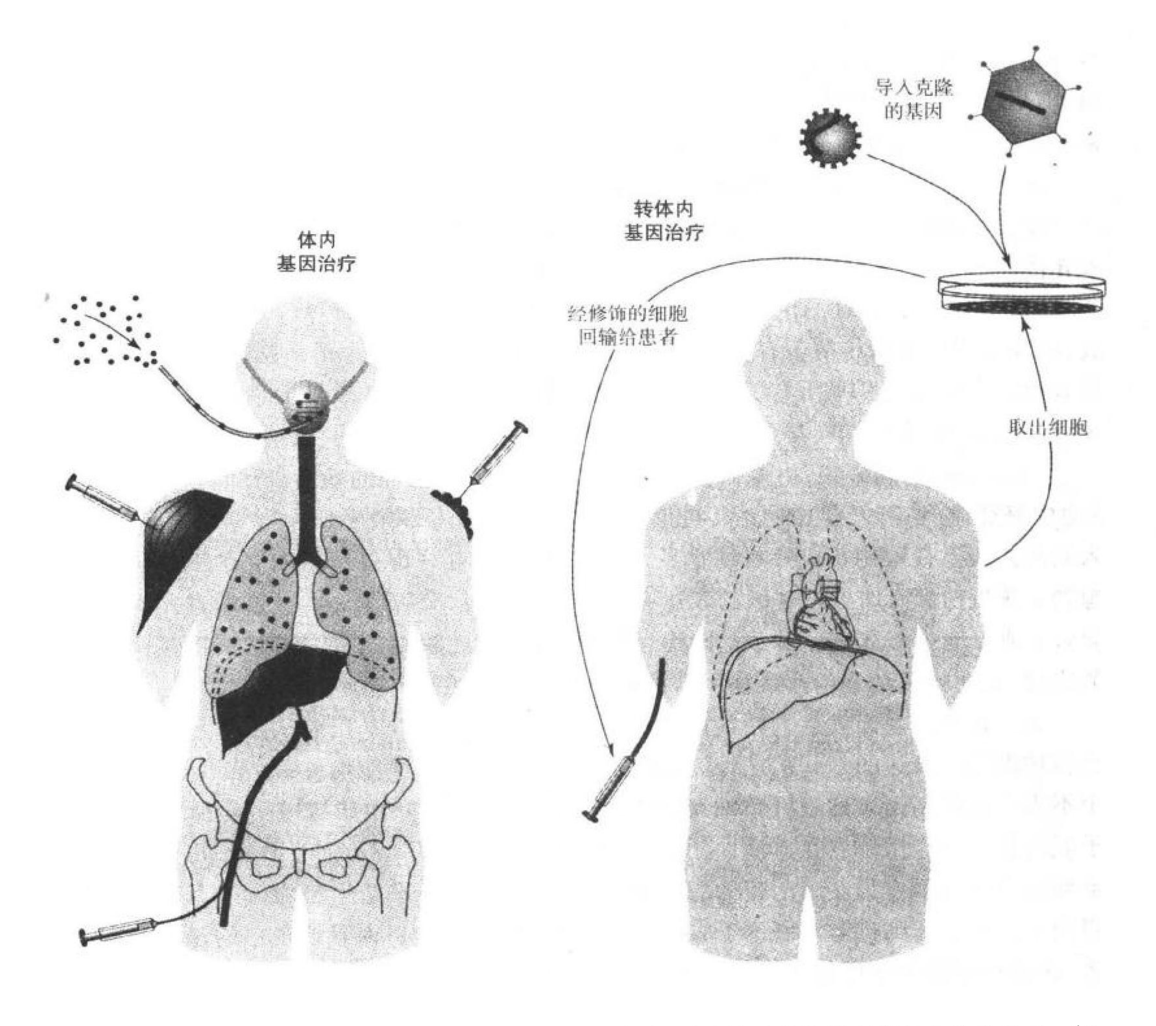


图13-5 运送基因治疗载体的常用方法。体内基因治疗中，治疗性遗传物质直接导入患者体内。例如DNA直接注射入肌肉或肿瘤内，经直接血管内扩散将DNA运送入肝脏或其他器官，或者作为气雾剂将载体运送至肺。在转体内基因治疗中，首先从患者体内取出靶细胞，在组织培养皿内暴露于基因治疗载体，所需的细胞群在经选择性扩增或筛选后，通常采用血管内扩散将已修饰细胞回输到患者体内。

因治疗是指从患者身体取出的细胞在组织培养皿内暴露于载体,然后将已摄取所需 DNA 的细胞再回输给患者。该方法具有可以精确控制暴露于基因治疗载体的细胞类型的优点,还易于使细胞暴露于高浓度载体以及其他凡是可能使这一过程更为完善的必需的条件。

体内基因治疗是指经血管内输注或直接注射到特定部位和组织的方法将载体直接导入患者体内。该方法通常更为简单,技术要求少,还能将载体运至体外操作不易达到的部位。体内基因治疗的缺点是载体暴露的条件难于控制,并存在重组序列广泛播散的可能,包括对生殖系的意外污染。

目前在本章中,我们已经讨论了能用来介导外源 DNA 进入靶细胞的技术,另一重要的问题是选择哪个基因进入何种类型的靶细胞。图 13-6 说明正在研究的一般策略,导入的 DNA 能够编码一种蛋白质以替代缺陷因子或通过正常宿主基因或新的外源基因的异常表达来达到治疗效果。这个基因可在细胞内或在较远部位局部性地发挥作用。

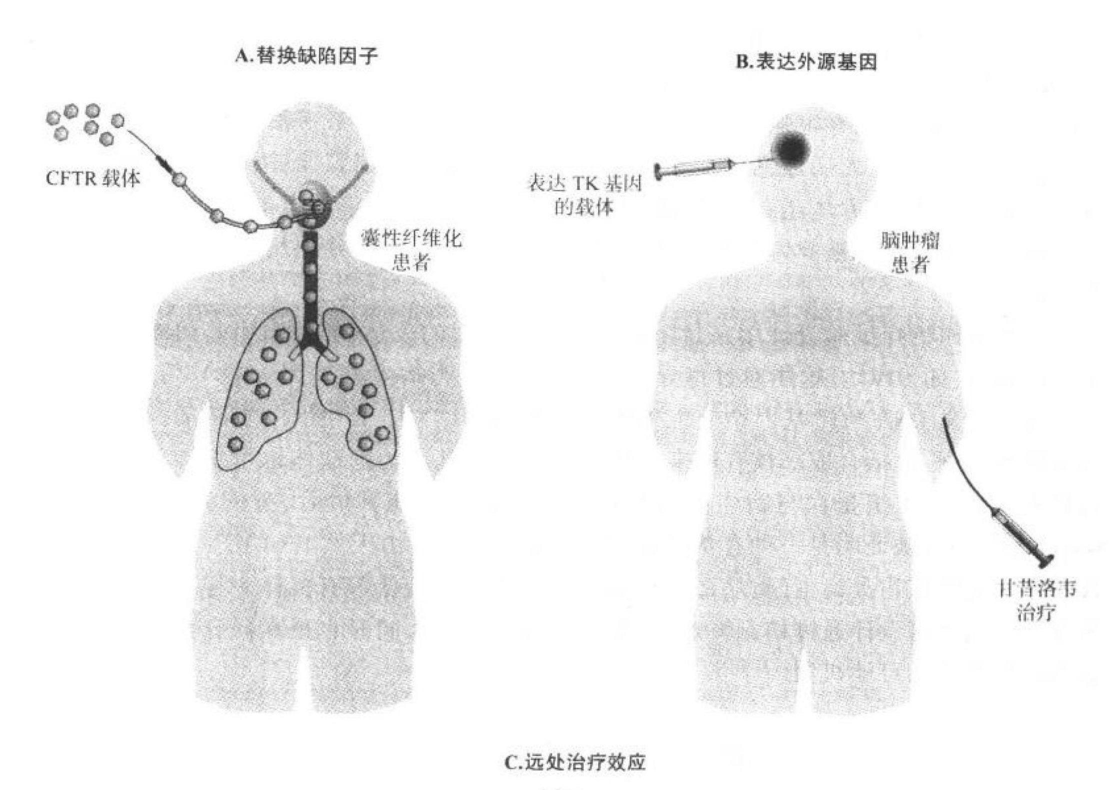
也许最显著且直接的策略是替代缺失的 DNA 序列以治疗隐性遗传病,这个实验性基因治疗试验的例子是应用腺病毒载体将正常的囊性纤维化跨膜调节因子(CFTR)序列导入囊性纤维化(CF)患者的肺内。理想中,这一方法能够完全替代关键类型细胞中的缺失基因,以达到完全治愈遗传病。对于像 CF 或肌萎缩这一类遗传病,这是一个严峻的挑战,因为对于这些遗传病,治疗基因必须以特定的细胞群体作为靶细胞。此外,还必须能高效地替代所有或几乎所有的关键类型细胞中的基因。大多数显性遗传病可能也很难应用基因治疗,特别是在涉及显性负调节机制或突变蛋白质获得功能效应时。

对许多遗传病来说,由基因治疗矫正缺陷还必须有精确的基因表达转录调控,例如:β地中海贫血(见第六章)的治疗,将正常的重组β珠蛋白基因序列导入骨髓细胞,需要导入基因的表达仅限制于成熟红细胞系,并能接近正常β珠蛋白基因的高水平表达。你会想到α珠蛋白和β珠蛋白基因的表达必须仔细地相互协调,而α或β链的过剩则会分别导致β或α地中海贫血。对于基因治疗领域来说,如此精细的基因表达调控是一个严峻的挑战,要达到成功地治疗疾病所需的调控程度无疑还需若干年的努力。

虽然在体内基因表达的精细调控还不可行,但对于治疗基因表达水平范围广泛并能在多种组织内表达的一些疾病,基因治疗可能是有益的。甲型血友病(见第七章)就是一个不需对重组基因表达进行精细调控的人类遗传病实例。由于这些患者所缺乏的Ⅷ因子凝血蛋白的作用位点在血液中,只要能将Ⅷ因子直接分泌入血液循环,几乎所有可接触到的组织均可作为基因治疗的潜在靶目标。此外,对于严重甲型血友病患者只要血浆Ⅷ因子水平提高到正常水平的1%就会有显著的临床改善。而且该蛋白的过表达未必有害,在某些情况下还可见2~3倍正常水平表达也无明显的不良后果。

除以正常序列替代患者的遗传缺陷外,某些基因治疗方法依赖于正常或外源基因的异常表达。例如,治疗性生长因子之类的正常序列过表达对于一些疾病的调整具有潜在的价值。DNA 疫苗领域或许是最有前景的基因治疗应用之一,肌肉内注射 DNA 指导一种危险的抗原表达,如微生物病原体或者甚至是一个肿瘤细胞,可证明是在某些条件下发生免疫化的一种很有效的方法。

有些基因治疗策略指导靶细胞中毒素的表达,最为常见的是肿瘤细胞,这种毒素可



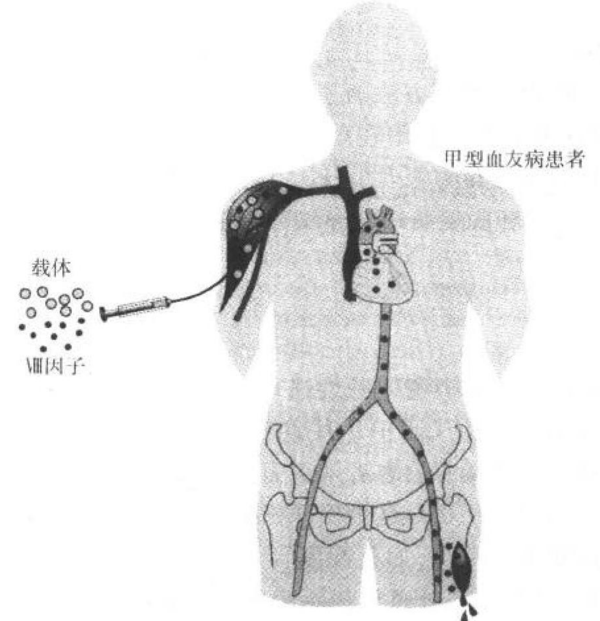


图 13-6 基因治疗的一般策略。基因治疗能够替换患者体内有遗传性缺陷的缺失因子，如将 *CFTR* 基因经由喷雾剂吸入囊性纤维化患者的肺内(A)。基因治疗也能使宿主基因以新的方式表达或使外源基因表达来达到治疗目的。在(B)中，胸苷激酶(TK)使细胞对随后的药物甘昔洛韦治疗敏感，该方法应用于脑瘤的治疗，则药物可杀死脑瘤细胞，而不影响正常的宿主细胞。例子(A)和(B)都依赖于给药位点处基因治疗载体的局部作用。(C)说明治疗基因产物远距离发挥作用。在此例中，基因治疗载体导入甲型血友病患者臂部肌肉内，产生的Ⅷ因子经血循环来矫正凝血障碍，治疗远处受伤部位的严重出血。

直接杀死肿瘤细胞或作为一种"药物前体"使肿瘤细胞对随后的治疗敏感。例如,由单纯疱疹病毒基因组编码的胸苷激酶(TK),能将通常无毒的药物甘昔洛韦(ganciclovir)转化成一种对哺乳动物细胞有很强毒性的复合物。这就是应用甘昔洛韦治疗疱疹病毒感染的基础,它能有效地杀死病毒感染细胞。将 TK 基因导入肿瘤细胞能够使细胞对随后的甘昔洛韦治疗敏感,该方法还有通过"旁观者"效应来杀死靶细胞附近的肿瘤细胞的优点,尽管这些细胞自身未被载体转导。这一机制可增强只进入部分肿瘤细胞的基因治疗载体的效能。

某些基因治疗试验设计用来封闭靶细胞有害基因的表达,如肿瘤癌基因或病毒组装必需的基因(如 HIV)。这能通过指导合成反义 DNA(与宿主靶基因 mRNA 互补的 DNA)来实现。反义 DNA 与靶基因杂交能够阻断其转录、RNA 加工及蛋白质翻译。反义策略也能由短序列,包括合成寡核苷酸来完成。还有一种基因治疗载体也可指导一种重组蛋白的表达,这种重组蛋白与宿主蛋白相作用并阻断其功能。最后,有一类称为**核酶**(ribozyme)的 RNA 被证明是一种有益的基因治疗工具。这些由天然 RNA 衍生而来的分子已被鉴定为某些植物类病毒(类病毒颗粒)的成分,和某些较低等有机体一样能够自我拼剪内含子。核酶可设计通过切割特定序列使靶 mRNA 失活,同时也正在研究其催化精确修正致病点突变的可能机制。

四、人类基因治疗实验

针对人类基因治疗的前景,美国国立卫生院(NIH)的重组 DNA 咨询委员会(RAC)创建了一个基因治疗分会以讨论由这种新方法所引起的科学和伦理问题。1989 年第一次审批研究会议上,RAC 认真评审了所有提出在美国实施的人类基因治疗实验,直到 1996 年这一职责大部分被转交给食品与药物管理局(FDA)。第一例批准的人类基因治疗试验始于 1990 年,是治疗两个由于腺苷脱氨酶缺陷导致的严重联合免疫缺陷(SCID)患者。

(一) 转体内基因治疗

腺苷脱氨酶(ADA)缺陷是一种常染色体隐性遗传病,由 20 号染色体上的 ADA 基因失活突变所引起。腺苷脱氨酶是嘌呤代谢中一种重要的酶,它的缺陷导致代谢物,特别是 T 淋巴细胞中毒物的累积。严重受累的患者几乎完全丧失 T 细胞,导致严重的免疫缺陷,如果不予治疗,通常在儿童早期死亡。这个疾病引起了特别的关注,并因以下几点原因将其作为基因治疗的候选对象。首先,如有配型合适的骨髓供者,ADA 缺陷患者能够通过骨髓移植而治疗。此外,简单的输血通常就能使患者病情明显好转。这些结果表明替换骨髓内的酶类足以达到治疗目的,甚至低水平的酶活性即可明显改善临床病情。骨髓细胞亚群已成功地转导了基因治疗载体并表达 ADA,这使得患者原来的 ADA 缺陷细胞也具有选择性生存的优势,导致由已纠正的细胞在骨髓中逐渐重建细胞群。最后,因为没有合适的骨髓供体而预后危险的患者则使这种新试验治疗的潜在危险合理化。

这个最初试验是转体内基因治疗的一个实例。从两个 ADA 缺陷的患儿收集的 T 细

胞在组织培养皿内经用指导 ADA 表达的反转录病毒载体处理,以此方法最早治疗的两个患儿现已随访了五年多。T 细胞数量及几项免疫功能的测定均已显著提高,治疗结束后的几年内均能检测到整合的载体 DNA 和 ADA 基因的表达。

尽管这些结果令人鼓舞,但由于缺乏合适的对照组比较以及同时还采用纯化的修饰型 ADA 酶治疗这些患者(对多数患者有效的标准疗法),使对治疗结果的解释复杂化。因此,直接的治疗效果是否归功于基因转移本身仍不清楚。

已在第七章详细讨论过的家族性高胆固醇血症,是下一个应用转体内基因治疗方法的遗传病。这些患者明显缩短的生存期正是进行这一实验的理由。首先外科手术切取部分肝脏,从切除的肝脏制备肝细胞在组织培养中生长以诱导细胞分裂,使之经反转录病毒基因治疗载体感染后 DNA 能稳定地插入基因组。病毒经改造能表达正常 LDL 受体基因。基因转移后,转导的肝细胞由门静脉再输入患者体内。虽然首例报道的患者体内 LDL 胆固醇有所下降,但由于该患者体内的 LDL 受体基因部分性质缺陷,可使患者对与基因转移无关的其他治疗方法起反应而达到治疗效果,致使对该项实验的解释复杂化.

(二) 体内基因治疗

第一个体内基因治疗实验是转移的恶性黑色素瘤的治疗。皮肤肿瘤被直接注射设计用以表达 HLA Ⅰ类基因 *B7* 的质粒 DNA, 它能与患者的 HLA 发生错配(见第九章 HLA 内容)。预计这个外源基因导入肿瘤细胞能够刺激患者自身免疫系统的免疫识别和排斥。考虑到非病毒治疗表现的较低风险和令人鼓舞的初步结果,这一策略正推广到其他类型肿瘤。

为了评价各种体内基因治疗方法,我们广泛应用动物模型研究了一些治疗基因表达水平明显提高和延长并具有生物学效应的重要例子。患血友病的狗在经腺病毒载体及其他载体,包括反转录病毒、AAV 及非病毒载体表达的各种蛋白产物治疗后,其体内的凝血因子Ⅸ水平明显升高。然而人体内试验的例数仍然较少。

第一例采用体内基因治疗方法的人类遗传病是囊性纤维化。现在已有一些患者把表达正常 *CFTR* 基因(见第九章)的腺病毒载体以喷雾剂形式吸入到肺内,然而至今结果仍令人失望,没有可检测的治疗效应的明显迹象。现在还有其他一些遗传病已开始体内基因治疗试验,在今后的几年内可望出现大量新信息及技术方法上的改进。

五、肿瘤的基因治疗

考虑到基因治疗方法产生无法预料的遗传改变的潜在危险,尤其是反转录病毒介导的插入突变的致癌效应的可能性,故基因治疗实验一般限定于那些患有严重危及生命的疾病及生存期短的患者。除上述的将 HLA *B7* 基因导入黑色素瘤细胞外,其他一些策略也被用于调节对肿瘤的免疫反应,包括导入各种不同类型的生长因子。人们也已尝试干扰突变癌基因的功能或替代如像 *P53* 抑癌基因的缺失。还有一些实验已尝试将“自杀基因”导入肿瘤细胞以直接杀死细胞或使其对随后的药物和放射治疗敏感。尽管少数研究

已显示诱导出所希望的免疫反应的令人振奋的迹象,但至今仍没有明确记载的对基因治疗方法产生反应的临床重要肿瘤实例,所以这些技术仍局限于实验领域。

六、基因治疗的风险和益处

图 13-7 总结了自 1997 年 6 月以来已被 RAC 和 FDA 提交评审的基因治疗方案类型。这些研究中绝大多数是治疗那些无法治疗或对标准的治疗无效的各类肿瘤患者。还有一些特定的遗传病也是治疗目标,尽管这样,最初的治疗试验也主要限于那些因无有效的替代治疗手段而预后不良的患者。

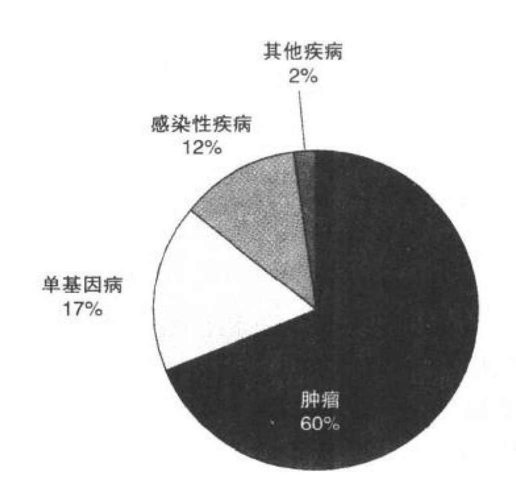


图 13-7 提交给 RAC 或 FDA 的人类基因治疗方案。至 1997 年 6 月共提交了 161 种基因治疗方法,其中 110 例用于肿瘤治疗,74 例肿瘤试验同时采用了免疫疗法,其中大约半数应用体内基因治疗。22 例肿瘤病例应用 TK 基因使肿瘤细胞对甘昔洛韦敏感,仅 28 例是用于单基因遗传病,其中半数以上是治疗 CF,19 例感染性疾病试验均是针对 AIDS,“其他疾病”一类中包括外周动脉疾病、类风湿性关节炎、动脉再狭窄和尺骨腕管狭窄综合征的治疗方案各一例。

到 1995 年 6 月,RAC 审批的共 106 种基因治疗方案治疗了 597 个患者。除一例 CF 患者用腺病毒载体治疗后引起一过性肺功能衰退外,还未发现与基因治疗载体或治疗基因本身直接相关的临床并发症的证据。令人放心的是,在这些最初试验中,尤其是与这些疾病的其他实验性治疗相比较,明显缺乏严重并发症或毒副作用。然而由于患者群体的性质,对晚期并发症的随诊有限。最可怕的并发症,即由反转录病毒插入所引起的肿瘤发生危险可能需要经过很多年才在临床上表现出来,因为由反转录病毒引起的基因改变可能仅是某种肿瘤发生中的“第一次打击”(见第十一章)。不过,现有的经验表明体细胞基因治疗一般是安全的且危险少。

那么益处又何在呢? 虽然我们已讨论过的一些基因治疗实验提供了对临床疗效极具启发性和鼓舞性的证据,但 100 多种审批的基因治疗手段中没有一种能够提供治疗效果的确切证据。虽然研究基因转移用于疾病治疗已近二十年,但我们认为基因治疗领域

仍处于初期阶段。

七、其他重组 DNA 基础上的治疗手段

尽管一般认为通过 DNA 技术生产的重组治疗蛋白不属于基因治疗范畴，但其对医疗实践的影响日益增长。表 13-2 列举了在现代分子遗传学到来之前根本不可能出现的一些新"药"。其中一些重组蛋白用于替代那些先前用标准生化技术从人或动物组织中提纯的天然蛋白质。

现在应用重组 DNA 方法生产的人胰岛素通常替代了从牛胰脏中提纯的牛胰岛素。其他一些重组蛋白也替代了那些以前从人类血浆或组织标本中提取的蛋白产品。重组蛋白不像从人组织中制备的蛋白产品那样总是要担心有病毒污染，它没有人类病原性病毒污染的风险。重组生长激素已替代了先前从人垂体中提纯的可能污染"慢病毒"的产品，这种"慢病毒"能引起严重的神经退行性疾病。20 世纪 80 年代早期 HIV 病毒污染血源，导致那段时间接受治疗的大多数严重血友病患者无意中输入了充满 HIV 的Ⅷ因子，致使大部分血友病患者死于 AIDS。这一悲剧是后来克隆Ⅷ因子基因及生产无病毒污染风险的Ⅷ因子蛋白产品的主要原因。重组蛋白工程也已用于生产疫苗，如乙型肝炎疫苗，以前是靠从人的血液中分离的病毒抗原。

表 13-2 重组 DNA 技术的制药产品

重组产品	靶疾病	较之现有药物的优越性
人胰岛素	糖尿病	比猪或牛胰岛素免疫原性更低
生长激素	生长激素缺乏	大量应用而无感染风险
重组Ⅷ因子	甲型血友病	无感染危险 较低的免疫原性
组织纤溶酶激活因子	急性冠脉血栓(心梗)或中风	与天然细菌产品链激酶相比更高效 较低的免疫原性
促红细胞生成素	贫血	NA
粒细胞集落刺激因子(G-CSF)	化疗后中性粒细胞减少症	无感染风险
乙型肝炎疫苗	预防乙型肝炎	无感染风险
α 干扰素	毛状细胞白血病、慢性肝炎	无感染风险
β 干扰素	多发性硬化症	NA
γ 干扰素	慢性肉芽肿病人的感染	NA

NA：以前无类似产品

基因重组技术使新型的生物药物成为可能，否则就不能满足大量的需要。例如，包括某些治疗作用的生长因子，像促红细胞生成素，能刺激红细胞的产生。治疗肾衰引起的贫血，粒细胞集落刺激因子(G-CSF)能够对抗肿瘤化疗带来的常见并发症，即白细胞数量骤降。

治疗用重组蛋白通常由细菌、酵母或组织培养的哺乳动物细胞生产。绵羊、山羊等家畜的转基因技术也用于生产表达重组蛋白。采用这种方法，重组蛋白直接进入乳汁，从而可能进行经济而高效地大量生产。

八、未来的趋势

虽然在过去的几年内，人们对基因治疗的兴趣显著增加，由这一领域发展起来的技术可能最终会与其他方面的医疗手段相融合。重组 DNA 技术的手段并非建立一个崭新的、孤立的领域，而是简单地扩展现有医疗的潜在领域。例如，包括目前一些治疗 ADA 缺陷的试验(见上面)的许多基因治疗方法，是依赖于将靶基因在体外导入骨髓细胞再通过标准的骨髓移植将这些骨髓细胞回输给患者。这一方法可视为现有的骨髓移植方法一种技术上的提高。事实上，所有的异体骨髓移植(将另一个体的骨髓导入患者体内)都将看作是一次完整的基因治疗，它是以携带致病基因正常拷贝的另一个体的骨髓细胞替换患者骨髓细胞的整个基因组。与之相似，通过基因治疗诱导肿瘤细胞内的潜在免疫原性基因表达，包括向肿瘤内直接注射免疫原性细菌产物，与传统的免疫疗法在概念上密切相关。

最后，基因治疗的一个有前景的新发展方向是生物反应器的创造，可将其植入患者体内长期生成治疗性蛋白。所设计表达重组蛋白的细胞能在植入装置内生长，而营养物质可进入其中以支持细胞生长，同时重组治疗蛋白也能扩散到血液循环，又能保护细胞不受宿主的免疫破坏。这样的植入装置广泛地应用于治疗许多蛋白缺陷性疾病，这将比其他形式的基因治疗更为安全。这一方法在概念上相似于常规意义上的治疗蛋白或药物扩散到患者体内。

基因治疗及其他相关的 DNA 重组技术在人类疾病的治疗中具有广阔的应用前景。改进治疗性 DNA 序列的传递和基因表达的有效调节技术将为许多遗传性和后天获得性疾病，包括常见多因素疾病，如动脉硬化、糖尿病、高血压及肿瘤的治疗，创造令人激动的新方法。随着这一疾病群的相关易感遗传因素的确定，就可能在出现症状之前鉴定高危易感人群，并开始予以预防性的基因治疗。重组 DNA 技术最终使精细矫正实际位点上目的基因的特定遗传缺陷成为可能，而在基因组内的另外部位无其他副作用的风险。伴随基因治疗及其他领域的持续发展，分子遗传学实验技术对人类疾病诊断的强烈冲击会使我们很快地认识到它在疾病治疗及预防领域中的辉煌应用前景。

推 荐 读 物

Friedmann t. A brief history of gene therapy. Nat Genet 1992;93 ~ 98. *A historic perspective on gene therapy, beginning with the early cell transformation studies in the* 1960*s*

Blaese RM, Culver KW, Miller AD, et al T lymphocyte -directed gene therapy for ADA-SCID: initial trial results after 4years. Science 1995;270:475 ~ 480. *Follow -up report on the first approved clinical gene therapy study*.

Wivel NA, Walters L. Germ-line gene modification and disease prevention: some medical and ethical perspectives. Science 1993;262:533 ~ 538. *A discussion of the issues surrounding germ-line gene therapy from the directors of the Recombinant DNA Advisory Committee* (*RAC*).

Crystal RG. Transfer of genes to humans: early lessons and obstacles to success. Science 1995;270: 404 ~ 410. *A general overview of the gene therapy field*

Rosenfeld MA, Collins FS. Gene therapy for cystic fibrosis . Chest 1996;109: 241 ~ 252. *A review of current progress and the major obstacles to gene therapy for cystic fibrosis .*

Roth JA, Cristiano RJ. Gene therapy for cancers: what have we done and where are we going? J Natl Cancer Inst 1997;89:21 ~ 39. *A detailed summary of the current status of cancer gene therapy, including trails completed or in progress worldwide.*

Orki Sh, Motulsky Ag. Report and recommendations of the panel to assess the NIH investment in research on gene therapy. 1996. *The report of an expert panel pointed by the Director of the NIH to review the status of gene therapy. The report can be obtained through the internet at the home page for the Office of Recombinant DNA Activities (ORDA) at* ***http://www.nib.gov/od/orda/index.html****. This website is also a useful resourse for additoinal up-to-date information about gene therapy, such as the complete list of approved gene therapy protocols and their current status, including the number of patients enrolled and comlications observed.*

第十四章　伦理的思考

聪明反被聪明误。

——Sophocles

医学遗传学,曾是较少数专家们涉猎的领域,现在由于其在遗传病病因鉴别方面的显著进展,正在进入临床医学的主流。但是无论在临床医疗方面还是在临床研究方面,遗传医学经常向患者、研究课题、家庭、医生、研究人员以及社会提出一些困难的选择。其中一些与其他医学学科所面临的窘境没有什么不同——现实的选择可能少于原来设想,赖以决策的信息可能残缺片面,因此迫使医生和患者在可能是最好的方法中选择一条向前的途径,是否成功并无把握。在医学遗传学领域,另外的复杂性可能出现:遗传信息是特别重要的,且其伦理内涵意义尤为深远,因为它可以预测目前健康的个体在未来可能患有的疾病,并且对家族中的其他成员(已出生或未出生的)可能有重要的影响。

纵观全书,读者或许有时为某些呈现的临床状况的伦理问题所困扰。结束本书最后一章的目的是概述某些伦理的一般原则,以便于用更清晰的术语阐述这种挑战性的状况。通过解说病例,概述医学伦理学的四个主要原则并且将其应用于遗传学。然而,读者不应期望对复杂问题作简单的回答;当其中两个原则冲突时常引起医学遗传学中的一些最困难的争议。然而,在许多实际临床问题中将会看到明显的伦理冲突部分是由于患者、医护人员和家庭之间交流不当的结果。

思考这些原则之前,简要地询问它们所依据的道德观点是恰当的。什么决定一次行为的益处?执行者是否尽了他或她的职责和义务?此人的动机是否道德高尚?或者简单地说行为的后果是什么?三个传统的伦理理论(责任理论、道德伦理学以及逻辑推理)分别代表那三种观点。其他当代理论超越个体范畴包含社区和更广泛的社会背景。充分思考医学伦理学的理论基础远远超出本章范畴,有兴趣的读者请查阅本章结尾的权威性参考文献。

一、尊重自主权

思考的第一个原则是尊重个人的自主权。自主权是自决,自控,或自治的同义语。尊重自主权基于这样一种主张即个人能够推理、决策、选择并且行动,因此有权管理掌握自己的命运。这一原则要求与患者探讨某种特殊措施(可能是一个方法或者一次遗传试验)的医生有责任提供足够的遗传信息,使患者对于他或她是否愿意接受这种措施作出有信息根据的、独立的、自愿的判断。这就要求解释它的益处、风险并探究可能出现的结

果。在医学遗传学环境下，此过程可分成三步：(a)预试验教育和咨询，即解释试验的风险和益处；(b)如果患者愿意，进行试验本身；(c)关于试验结果的实验后咨询。遗传咨询必须是启发式的原则直接产生于对自主权的的尊重，这已在第十三章讨论过。

病例 1——随访选择困难的遗传试验

一位 36 岁妇女有明显的乳腺癌和卵巢癌家族史。了解到 *BRCA1* 基因(见第十一章)检测的有效性后，她寻求一项研究方案，发现自己是合适的候选人。经过培训，她知道了她的 *BRCA1* 基因突变状况的风险和益处，她选择进行试验。她的姐姐(见图 14-1)患有双侧乳腺癌，其 DNA 样本显示外显子 11 处有 2 个碱基缺失。因为这是移码突变，高度提示它是引起疾病的突变。患者再次确认她渴望了解自身状况，试验继续进行，结果发现她也携带这种突变。她深深地关注这一试验结果，并且不相信监控策略会防止她患癌症。经随访选择的广泛讨论之后，她选择预防性双侧乳房切除术以及整形重建手术。一年后，当决定不再生育时，她切除了卵巢。

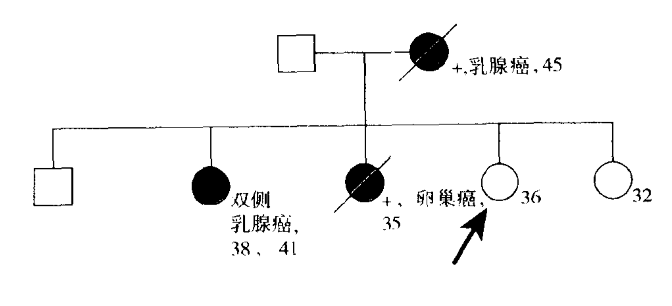


图 14-1

她的妹妹(见图 14-1)也有 50%风险携带同样 *BRCA1* 基因突变，家人告诉她关于这项研究方案检测的有效性后，她有点勉强地同意去诊所。预试验期间，她表明在二十几岁时有高度风险行为的背景并且描述了等待 HIV 试验结果(最终报告为阴性)的特别可怕经历。目前她已经 32 岁，自谋职业，她也担心一旦发现她是 *BRCA1* 基因突变阳性会因此失去健康保险。她选择不做检测。家族中的其他成员，尤其是先证者深深地关注她的决定，并且对医生施加巨大压力，进行研究探讨“开导开导她”。

评论

频繁乳房 X 线照射加强对 *BRCA1* 突变携带者的监控是否会改善生存，目前对这一信息的了解并不充分。现在对卵巢癌的早期检测尚无有效的监控。正如某些人仍会由术后残留的少量上皮组织发展成致命性恶性肿瘤一样，先证者选择的外科手术并没有 100%把握除掉癌症风险。然而，本病例中经适当的培训、检测、咨询措施，先证者让她选择最适合的方法来行使自主权，并且这些决定受到她的主治医生们的支持。

尽管她们的遗传风险是相同的，她的妹妹却作出了不同的选择。她的逻辑来自她个人观念和生活经历，完全有情可原。尊重自主权需要医生同样认真考虑她的决定就像对

待她的姐姐一样，并且不参与其他家族成员改变她的想法的工作。

病例 2——儿童检测

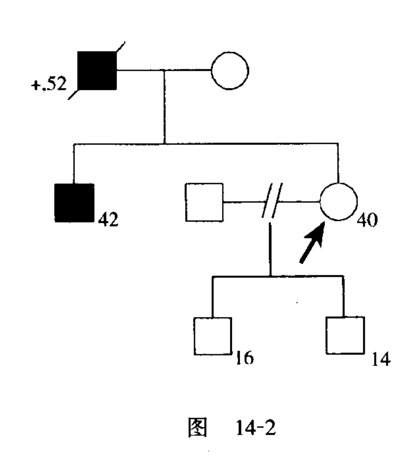


图 14-2

一位 40 岁妇女，其父亲死于 Huntington 病（HD）（见第九章），她的兄弟现在正患此病，她希望了解自身携带 HD 突变的风险（图 14-2）。经过试验前广泛咨询后，她决定为自身未来和保持心中平静，她愿意了解自身状况。试验结果阳性给本来已不稳定的婚姻添加了额外的压力。她丈夫起诉离婚，留给她很有限的抚养费照顾两个儿子。6 个月后，她返回诊所要求对她的 14 岁和 16 岁两个儿子进行 HD 检测。询问其理由时，她表示经济窘迫使她不能送两个儿子都上大学，她宁愿根据两个儿子中哪一个更可能有正常寿命期限而作出决定。

评论

这种状况提出一个问题，即代表孩子们作出决定的双亲自主权的限制。许多常规医疗机构认为这是理所当然；一般说来双亲能够授权或者拒绝同意对他们的孩子进行任何医疗措施，除非此项措施具有潜在威胁生命的风险（例如，拒绝给脑膜炎儿童服用抗生素）。然而，就这一病例而言，HD 突变检测会为这些孩子的未来提供具有深远影响后果的信息。儿童期检测剥夺了孩子们对未来机会的选择，甚至他们中的一个或者全部都有可能决定当他们长大成人时，他们宁愿没有这样的信息（正如大多数有 HD 风险个体目前决定的那样）。因此，母亲对检测信息的渴望不得不被孩子们的最佳利益权衡。过去几年里关于这个问题有大量的争议，公众舆论认为不应该对孩子们检测成人期发作的疾病；除非有特殊有效措施而且必须在 18 岁之前开始。这就保护了对孩子们未来自主权的尊重，而没有一般地带来任何附加的医疗风险。然而，任何有关双亲自主权的侵犯必须慎重实施。

新生儿筛查（见第十二章）被视为一种特例，在这种特殊情况中，有一种严重和不可逆疾病（如苯丙酮尿症）要求立即实施有效治疗。检测的意义是无可争辩的以至于在大多数州，这样的筛查是强制性的且常在没有征得双亲同意的情况下实施。理由是这种善行的重要性（如下所述）超出了双亲自主权的范畴。然而，尽管能提出有说服力的论据，但即使在这种情况下也应该尽最大努力征得双亲知情同意，因为这提供一次有益的教育机会。

病例 3——轻度障碍终止妊娠

一对职业夫妇来遗传诊所要求对他们的 6 岁女儿进行检查，其女儿已诊断为轻度智

力障碍(图 14-3)。复查检验记录表明她对教育活动反应很好,她的智商是98,有充分理由表示她能发挥正常水平的学习功能。检查任何特殊原因都是否定结果。完成评估时,双亲询问再生一个患有类似智力障碍孩子的再发风险。尽管现在没有直接有效的检测方法,经验证据表明类似智力障碍的再发风险可能大约为 7% ~ 8%。双亲把女儿的智力障碍看得很重,尤其与他们接触的社会圈子中同龄儿童比较,表明这种风险似乎较高。随后母亲表示她已妊娠 8 周。她声明在已有一个女儿之后,再生一个患有这种智力障碍的孩子对她来说是太大的负担,她需要求助于终止妊娠。

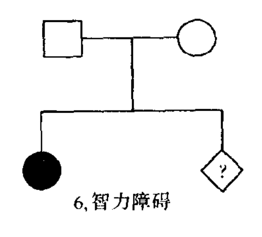


图 14-3

评论

这对夫妇将相对低风险的轻度智力障碍看得非常严重,每一位医学遗传学家都遇到过类似这样的病例。尊重自主权原则表明是启发式咨询原则。意在保护这对夫妇的权利,即根据他们的经验,作出最适合的结论。在一个无需医学解释就允许终止妊娠的社会里,毫无疑问,这对夫妇能够找到一位愿意参与终止妊娠的医务人员。然而,特别麻烦的是,这对夫妇在其女儿面前讨论这些问题的方式,反映了一种潜在的有害环境。在一次随访避开他们的女儿与双亲讨论时,提出了这一问题。在最初反对之后这对夫妇同意进行随访咨询。这是在这种情况中,遗传学家的责任心超越了原来服务需求的例子。

病例 4——性别选择

产科医生提到一对夫妇有 4 个女儿(见图 14-4)。他们决定在第 4 个女儿出生后不再妊娠,现已多少采取一些限制手段。然而,出现了意外妊娠,他们对继续还是终止妊娠争论不休,他们决策中的主要问题是希望有一个男孩。他们要求产前诊断确定胎儿性别,如果胎儿是女性,他们计划终止妊娠;如果胎儿是男性则继续妊娠。

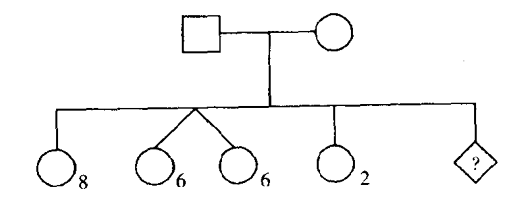


图 14-4

评论

利用产前诊断选择性别引发了许多令人棘手的问题。一方面,在美国要求流产有效使得这对夫妇的要求变得完全合法。一些专家把尊重自主权理解为应该支持双亲这种选择。但是女性性别不是疾病;它只是性状。普遍承认性别选择产生许多棘手的道德后

果。在大规模情况下是非道德的事物在小规模条件下就能符合道德吗？最近，一个医学研究所的专家小组裁决："利用产前诊断确定胎儿性别以及随后利用流产作为优先选择胎儿性别为目的是遗传服务的误用，这是不恰当的，同时也不应该受到卫生专业人员的鼓励。概括地说，人们不应该利用生殖遗传服务来达到优生目的，而应该着眼于在生殖选择上加强个体控制。"尽管强烈声明，临床遗传专家们按各自对性别选择的观点分类。近期调查表明美国有34%的遗传学家在上述情况下会进行产前诊断。在过去的20年里，这个数字一直在增加。

病例5——产前诊断选择患病儿童

一对夫妇都患有软骨发育不全，这是一种由成纤维细胞生长因子受体3(见第十二章)基因突变引起的显性侏儒症。他们来到遗传诊所要求产前诊断。他们担心胎儿可能患有纯合型软骨发育不全，这是一种致死表型(图14-5)。然而，在讨论过程中，他们表示他们强烈希望有一个像他们自己的孩子，患有(杂合性)软骨发育不全，而不是具有正常身高的纯合性个体。因此一旦胎儿是纯合正常的或纯合患病的，他们渴望终止妊娠；而如果胎儿患有杂合性软骨发育不全，则继续妊娠。

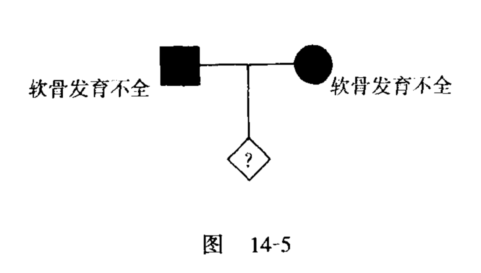


图 14-5

评论

这种情况可能并不经常出现，但却提出了极其艰难的挑战。帮助选择受累个体是否符合孩子的最佳利益？面临这种情况的许多遗传学家会拒绝参与这部分要求。故意恶化后代临床状况的决定似乎违背了对孩子有益的原则(如下所述)。如果进行特殊受累胎儿的选择，孩子甚至会起诉他的双亲？这样的案例被称为"错误的生命"，孩子采取法律行动指控双亲，认为他的生命是如此的充满痛苦，最好从未出生过。通常这些案例被法庭驳回，证据是与实际上根本没有生命相比，任何生命形式都是弥足珍贵的(这恰恰与称做"错误的出生"形成对照，"错误的出生"是应用于诉讼案件的术语，双亲通常以此指控医生，为他们没有受到生出畸形孩子风险的警告而辩解，如果他们了解这种信息，他们会设法避免的。)。

二、行善与无害

医学伦理学的这两个基本原则有时分开考虑，而实际上有许多共同之处。简而言之，行善就是受命将正在考虑的无论什么措施的益处最大化。无害是尽"首先无害"的责任，即风险最小化。在希波克拉底誓言中发现，几乎所有医学伦理学评论家都已经概述了思考这些基本原则的重要性。当与尊重自主权原则相结合时，维护那些原则要求医务人员充分描述风险和益处，以便使患者做出知情决定。这与推荐一个特殊的医疗过程更

为家长式的手段不同,即使这种手段可能是无危险而善意的。

病例 6——重症婴儿的高风险实验疗法

一个出生两天的婴儿患有严重的高血氨症。他的哥哥患有高血氨症,出生后三天死亡,随后高血氨症被诊断是鸟氨酸转氨甲酰酶缺陷(OTC)的结果,这是一种 X 连锁尿素循环紊乱,在出生后头几天几乎总是致命的(图 14-6)。检测表明母亲是携带者,第二次妊娠期间的产前诊断表明这个男孩将会受累。然而,由于双亲强烈反对流产,他们希望继续妊娠。但是婴儿现在病情严重。面对血氨升高,可用的常规疗法束手无策,正在考虑腹膜透析法。然而,恰好刚刚开发一种治疗 OTC 缺陷症的基因实验疗法,即把正常 *OTC* 基因拼接到腺病毒载体上,通过门静脉内注射大量的病毒接种物而进入肝脏。OTC 缺陷症动物模型的实验研究已经表明对尿素循环紊乱有短暂修复作用,这可能是帮助患儿度过急性危险期并逐渐过渡到维持疗法。到目前为止,这种新型基因疗法对人类的潜在毒性还未探索过。详细解释风险和益处,至少是尽可能的了解后,双亲选择进行基因疗法,并给予首次剂量。一天后,观察到血氨出现短暂的轻度降低,但几小时后婴儿死于暴发性脓毒血症。

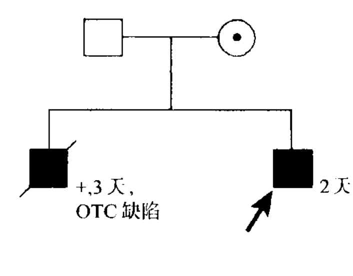


图 14-6

评论

新疗法具有未知的风险和益处,它的临床研究试验向知情同意过程提出了具有特殊意义的挑战。本例中,对患儿的潜在益处以及针对这种致死性疾病选择的先进疗法(行善)不得不权衡这种可能性,即由于不知道大量腺病毒可能引起的反应,这种侵袭过程可能引起伤害(无害)。在过去三十年里,对这些问题的广泛思考,已经导致在临床研究中控制使用人作为研究对象的原则的发展,其中包括向患者提供任何诸如此类实验治疗前,由法理学审查委员会(IRB)审查这些治疗方案。IRB 审慎考虑提议疗法的风险和益处,同时考虑更为传统的现行疗法,看此疗法是否是恰当的步骤。所有这些建议中一个至关重要组成部分是知情同意过程,此过程必须包括研究人员充分揭示风险和益处。

病例 7——不考虑终止妊娠时产前诊断的适当性

一对夫妇来到遗传诊所寻求囊性纤维化的产前诊断(见第九章)。他们的第一个孩子出生 5 天后死于胎粪样肠梗阻,这是一种囊性纤维化相对罕见的并发症,涉及小肠梗阻、穿孔和腹膜炎 (图 14-7)。当意识到目前怀孕也会有 25%的风险生出患儿时,他们渴望了解这种信息以便更好地计划未来。他们对囊性纤维化了解得很清楚,得知此病足以存活步入成年,他们决定不终止病态妊娠。

评论

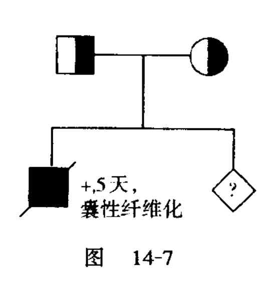


图 14-7

在这种情况下，尊重双亲自主权将会支持进行所需的产前诊断，向双亲提供需要计划未来的信息。然而，由于该信息不会改变结果，假如羊膜穿刺术有大约1/400的机会诱发流产，那么进行羊膜穿刺术是否合理？对胎儿而言，行善和无害可能引起争议即双亲的愿望应该被否决。但是因为流产的风险是如此之小，并且就父母的平和心境来说，其益处是明显的，所以通常尊重这类要求。

三、隐私和保密

关于第三者是否应该了解有关他们的医疗信息，一个普遍接受的原则是个人应该有权利作出知情决定。这当然包括遗传信息。第三者可能包括保险公司、雇主、配偶、家庭、教育机构、研究者以及其他人员。保密原则是指除了在显然易见的情况下，对医生-患者之间提供的信息进行保密。如果没有隐私和保密，医务人员和患者之间的信任将产生危机。然而，保密不是绝对的；例如，当患者向精神病医生描述一个即将发生的真实的谋杀计划时，医生有义务警告有生命危险的人。

病例8——发现错配父子关系

一对夫妇和他们的三个孩子来到遗传诊所咨询结肠癌，此病已累及父亲、父亲的姐姐和母亲，以及许多其他亲属(图14-8)。复查病志显示所有受累个体在确诊为癌症时已有上百个息肉，诊断为家族性息肉病几乎是确定无疑的(见第十一章)。用位于5号染色

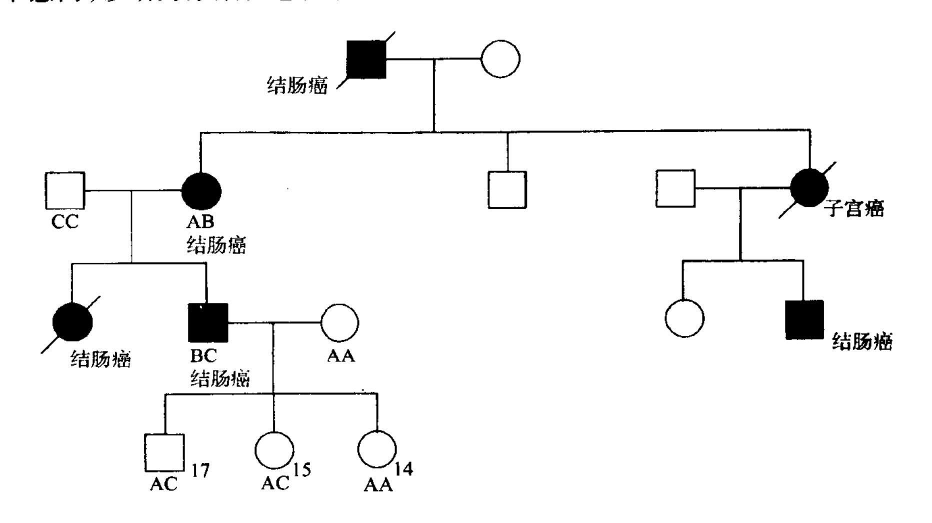


图 14-8

体长臂的 *APC* 基因探针进行连锁分析，发现致病基因与基因内标记物 B 等位基因连锁传递（见图 14-8）。因为治疗可能需要在 18 岁前开始，所以家族性息肉病是可以在儿童期进行检测的疾病之一。在完全知情同意的情况下，经孩子们的同意，以及广泛的家庭培训，用紧密连锁的标记物对三个孩子进行连锁检测。连锁分析确定两个稍大孩子从父亲那里获得了正常基因。然而，DNA 检测显示年幼女儿与其生物学父亲的 DNA 不匹配。若干其他多态标记分析证实有很大可能她的生身父亲是其他人。与其母亲进行的一次私人谈话中，她承认这种可能性即这个女儿是在很久前至今令她后悔事件中与另外的男人所生。她恳求诊所医生保密。因此，诊所人员只告诉这对夫妇和三个孩子，在这种情况下其父亲的家族性息肉性突变没有遗传给三个儿女中的任何一个。

评论

医学遗传诊所进行 DNA 分析时，发现父子关系不匹配并不是罕见的事件。事实上，作为检测过程存在的可能风险之一，在检测前应该无例外地提出出现这样结果的可能性。尽管某些人认为，父亲有权知道他女儿的生物学状况会迫使这种意想不到的信息泄密，但是，与母亲关系的秘密性以及她想保密的愿望通常权衡得重些。

病例 9——当警告义务与保密原则冲突时

一对夫妇来到诊所，因为他们首胎出生仅一周大的孩子已确诊为 Down 综合征（图 14-9）。体检显示有典型的特征，核型为 46，XX，－14，＋t(14;21)。双亲了解到该型 Down 综合征可能有潜在高度的再发风险（见第八章）。他们要求进行染色体研究。结果发现父亲携带 t(14;21)罗伯逊易位。已知其母亲有三次前三个月的流产史，因此不得不考虑这种易位是由遗传获得的可能性。在本例中，他的两个妹妹可能也是携带者，因此生出患有 Down 综合征孩子的风险增加。当提出对其双亲进行检测时，其父亲坚决拒绝讨论此事。家庭关系已经变得紧张。他因携带染色体重排导致生出 Down 综合征患儿而感到沉重的羞辱和有犯罪感。这位父亲决意不让家庭中的其他成员知道此信息。咨询员因此面临着困境，一方面有义务警告他两个将要分娩的妹妹，生出患有 Down 综合征孩子的风险增加，另一方面需要尊重医患关系的保密原则。

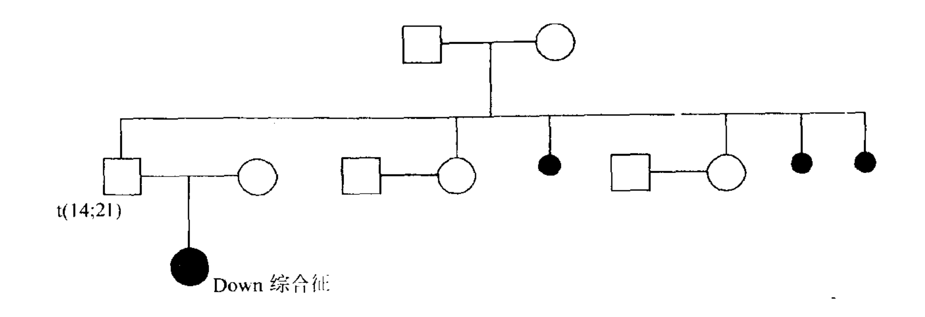


图 14-9

评论

在这种情况下,遗传咨询者的首要责任是充分解释清楚对亲属保密的后果。目的是帮助父亲考虑那些后果,认识到这样做的后果可能会比他所害怕的结果更糟糕。几乎在所有的例子中,由于有时间和机会进行理智的讨论,当对亲属有真正的风险时,家庭成员会认识到需要解密。在极少不能这样解决的情况下,目前还没有实质性的指导帮助医务人员确定这些神圣原则中哪一个更重要,而其他人患病风险的严重性应该是作出决策的主要决定因素。

四、公正和平等

公正和平等的原则是指在社会其余背景下如何对待一个人。利益分配公平吗? 负担分享了吗? 医疗,社会福利以及可能的有益研究是否局限于那些有经济来源或有明显社会价值的人们?

病例 10——遗传易感性有失去就业机会的威胁

一位无症状 25 岁妇女来遗传诊所检查,她有结肠癌和子宫内膜癌家族史。其家族史明显提示遗传性非息肉性结肠癌(HNPCC)(见第十一章)。作为研究方案的一部分,对其家庭进行评估,发现 *MSH2* 基因有一个突变。病人希望了解她的状况,广泛交代各种可能结果后,经检测发现她携带 *MSH2* 突变,这表明她一生中有 75%的风险可能发生结肠癌。然而,常规的结肠镜筛查可能降低这种风险。作为一名拥有工商管理学硕士学位的有抱负的商业专家,6 个月后,她在一家大公司应试一个有声望的职位,并开始试用。作为就业前健康评估的一部分,要求她填写一份详细表格,报告其结肠癌和子宫内膜癌家族史。关注到这种情况,公司和她的医生联系,尽管事实上检测只是研究方案的一部分,该医生复印并交给公司她的 *MSH2* 突变结果。看到未来疾病潜在的高发风险,人事主管撤消了对她的雇佣。

评论

在本例中,一位健康人由于对未来疾病的遗传易感性而被拒绝雇佣。这与通常的公正和平等原则不相容,这些原则认为所有的人应该有同等的就业机会,这是根据他们的信用而不是其他他们不能控制的信息。此外,她的医生违背了保密原则。直到近来,诸如此类的事例使很有希望的就业者处于无助状态,但是(正如第十章所述)平等就业机会委员会近期出台的条例认定雇主的行为违法,因此她很可能把公司诉诸法庭而重新恢复工作。

病例 11——由于遗传易感性失去健康保险

一位有明显乳腺癌和卵巢癌家族史的妇女登记进行旨在识别遗传病因的研究。经

过广泛的培训后，她表示愿意对 *BRCA1* 或 *BRCA2* 基因中可能的突变进行检测(见第十一章)，检测显示她的 *BRCA2* 基因存在无义突变，她母亲也有这个突变并因乳腺癌而濒临死亡。在 30 岁时，各种治疗方案的有效性不是完全清楚，但是在听取各种可能性后，她选择每隔 6 个月进行乳房 X 线照相检查，同时每月乳房自检和每隔 3 个月的职业健康检查。她的第一张乳房 X 线照片没有显示恶性肿瘤的证据，但是她的健康医疗保险商拒绝为检测付费，断然声称 40 岁以下的妇女没有必要进行乳房 X 线照相。她的主治医生试图代表她的利益而进行干涉，认为编造需要乳房 X 线照相的医疗病历可以帮助她，无意中泄露了研究已经显示她携带有 *BRCA2* 突变的事实。一星期后她收到健康保险商的一封信，通知她保险单已被取消。她是一个自谋职业的作家，不可能获得集体保险。考虑到她的有限财力，她得出结论乳房 X 线照相筛查不再是切实可行的，因此失去了医疗随访。

评论

毫无疑问这违反了保密原则。这一病例是否也可认为违反平等和公正原则取决于人们将医疗保健看作是权利还是特权。如果是特权，那么对个人的拒绝尽管对个人来说是巨大的不幸，但并不意味着违背伦理原则。另一方面，如果赞同这样一种观念即获得医疗保健是基本权利，则根据个人无法控制的易感性信息而遭拒绝在伦理学上是不被接受的。正如第十章所述，这种情况并不罕见，并随着更多的人了解其未来患病风险而变得更为普遍。从最严格意义上讲，我们都有预先存在的病情，即 DNA 的改变会使我们患某些疾病的风险增加。如果预先存在的病情是足以丧失健康保险的充分理由，在未来的十年或者二十年谁还会继续保险呢？迫切需要综合立法来否决利用遗传信息设立医疗健康保险费或拒绝承担保险。

五、结　论

本章无意详尽阐述应用于医学遗传学的伦理学的细节，也不想描述几百年来使得这四大原则普遍被接受的思想和学术成就。目前更为活跃的学术成就是在医学伦理学领域。有兴趣的读者可以在列出的参考文献中得到有助于他们理解的更详细的论述。正如我们在以上几个病例中提到的那样，人们普遍接受的是这些原则相互之间不完全一致的情况频繁出现。从这些例子中也可以看到，表面上出现的其实代表意义深远的伦理学困境的许多情况可以通过患者、家庭和医务人员之间深思熟虑的理智的对话得到解决。

这本医学遗传学教科书现在接近尾声。未来医生们面临的主要挑战是权衡使诊断更为准确、治疗更为有效的技术上巨大进展和与疾病中的困难做斗争或者受到疾病威胁时，保持温暖、同情、符合伦理的态度来对待患者和家属的需要。因此，未来的医生们要谨记俄罗斯谚语："让你的大脑保持爱，让你的心情明智。"无论现在还是将来，我们对我们自己和我们的病人都应该做到问心无愧。

推　荐　读　物

Andrews LB, Fullarton JE, Holtzman NA, Motulsk A, eds. Assessing Genetic Risks. Washington, DC: National

Academy Press, 1994. *The wide - ranging report of an expert panel of the Institute of Medicine, which considered many legal, ethical, and social consequences of advances in genetics.*

ASHG/ACMG Report: Points to consider: ethical, legal, and psychosocial implications of genetic testing in children and adolescents. Am J Hum Genet 1995;57:1233 ~ 1241. *A thoughtful review of the issues of testing in children, concluding that in general testing should be offered only if interventions are needed before age* 18.

Beauchamp TL, Childress JF. Principles of Biomedical Ethics. New York: Oxford University Press, 1994. *An authoritative reference on ethical theories and their consequences.*

Knoppers BM, Chadwick R. The Human Genome Project: under an international ethical microscope. Science. 1994; 265:2035 ~ 2036. *A scholarly review of the major ethical dilemmas in medical genetics, now being accelerated by the Human Genome Project.*

Wertz DC, Fletcher JC. Ethics and medical genetics in the United States: a national survey. Am J Med Genet 1988; 29:814 ~ 827. *A revealing survey of responses of* 295 *U.S. medical geneticists to* 14 *clinical problems and* 3 *screening situations that required a moral choice.*

词 汇 表

癌基因(oncogene)——原癌基因的一种改变形式(由于突变、过度表达、扩增引起),参与肿瘤的转化。

半保留复制(semiconservative replication)——用于 DNA,指每个子代 DNA 链都含有一条亲本链和一条新合成的链。

半显性(quasidominance)——具有相同隐性性状的纯合受累个体与杂合个体婚配产生的遗传模式,以致可在两代或连续数代出现纯合受累者。

包装细胞系(packaging cell line)——一种工程细胞系,它通过导入重组基因治疗载体缺乏的主要病毒基因来生产重组反转录病毒。

变形(deformation)——由于体内或体外的机械力作用于胎儿导致身体局部的异常体形、形状和位置。

变性梯度凝胶电泳(denaturing gradient gel electrophoresis,DGGE)——一种检测密切相关的 DNA 序列之间差异的方法。双链 DNA 样品在梯度递增的变性凝胶中解链。根据 DNA 序列内容的不同,双链沿着凝胶的不同位置解链,导致泳动的变化。

标记基因座(marker locus)——指能够容易地检测其等位基因的基因座,它可以是表达基因的一部分,也可以不是。

表现度(expressivity)——一个突变等位基因表型的性质和严重性反映了它的表现度。可变的表现度为常染色体显性性状的一个常见特征。

表型(phenotype)——基因型与环境因素相互作用后观察到的结果,是特定基因或一些基因可观察到的表现。

表型模拟(phenocopy)——一种由模拟遗传决定性状的环境因素产生的表型。

不等交换(unequal crossing over)——染色体上错误排列的相似 DNA 序列间的交换,导致核苷酸的缺失或重复。这种不等交换可以解释血红蛋白的若干变异体,例如 α 地中海贫血和 Lepore 血红蛋白。(见正文的解释)

不分离(nondisjunction)——指在减数分裂Ⅰ期两条同源染色体,或减数分裂Ⅱ期及有丝分裂时一条染色体的两个染色单体的不分开,从而都进入同一子细胞,而另一子细胞未接受该染色体。

不外显(nonpenetrance)——有突变基因型的个体缺乏相应基因表型的临床表现。

不一致性(discordant)——指一对孪生同胞中,其中之一表现出某种性状,而另一个不表现该性状。

蚕豆病(favism)——一些 G6PD 缺乏的个体因食用蚕豆引发的急性溶血性贫血。

插入(insertion)——一种染色体的结构异常,即一条染色体的部分物质插入一条非同源染色体。

查证(法)(ascertainment)——在遗传学研究中对所包含个体的选择。

常染色体(autosome)——除性染色体或线粒体染色体之外的任何染色体。

常染色体疾病(autosomal disease)——由 22 对常染色体中的一条染色体上的基因编码的疾病。

常染色质(euchromatin)——胰蛋白酶 G 显带中亮染的染色质,认为其中含有活性转录或具有转录潜能的基因。

沉默基因(silent gene)——无可探测表型效应的突变基因。

初级转录物(primary transcript)——一个基因的直接 RNA 转录物,包括内含子和外显子。

纯合子、纯合性(homozygote, homozygous)——指一对同源染色体上特定的基因座有一对相同等位基因的个体。

错义突变(missense mutation)——DNA 单碱基替换导致产生指定不同氨基酸的密码子。

单倍体(haploid)——正常配子的染色体数,只具有每对染色体中的一条。人类单倍体染色体数为 23。

单倍体同一(haploidentical)——在两个个体中,一个特定基因座共享一个单体型而第二个等位基因却不同。例如在 HLA 分型中,双亲之一与孩子总是单倍体同一,因为孩子继承双亲一方的一个 HLA 等位基因,从另一方继承了第二个等位基因。

单合子的(monozygotic)——指起源于一个合子或受精卵并在遗传学上完全一致的双胞胎。

单基因病(single-gene disorders)——对患者表型有较大影响的单个突变基因引起的疾病(如镰状细胞贫血)。

单链构象多态性(single-strand conformer polymorphism, SSCP)——根据单链 DNA 构象的改变导致它们在凝胶中泳动的变化,检测序列间单个核苷酸差异的一种方法。

单亲双体型(uniparental disomy)——一条特异染色体的两个拷贝都遗传自一个亲体,而不遗传另一亲体的染色体拷贝。

单体型(haplotype)——指一条染色体紧密相连的两个或两个以上基因座一组等位基因的基因型,通常作为一个单位遗传。例如 HLA 复合体。

单体性(monosomy)——指一对染色体中丢失一条染色体的疾病,例如 Turner 综合征中为 45,X。

蛋白质印迹(Western blot)——与 DNA 印迹类似的技术,以检测蛋白质。通常运用免疫学方法。

蛋白质自杀机制(protein suicide mechanism)——在显性疾病中,一个突变的亚单位可导致整个多聚蛋白质功能的丧失,例如,胶原。

倒位(inversion)——染色体的一种结构异常,其中染色体节段颠倒后重接;若倒位的节段包括着丝粒,则是臂间倒位;否则是臂内倒位。

等臂染色体(isochromosome)——一种异常的染色体,染色体的一臂复制(以致由相同位点

以相反序列形成等长的两臂），而另一臂缺失。

等位基因（allele）——在一既定基因座上一个基因的替换形式。

等位基因特异性寡核苷酸（allele-specific oligonucleotide，ASO）——设计合成的寡核苷酸，可在适当条件下与特异序列杂交而不与其相关的序列杂交。用针对每个等位基因序列设计的 ASO 甚至可容易地检出单个核苷酸的变异。在几种设计相似、用来区分密切相关等位基因的方法中，ASO 还可用作 PCR 引物。

等位基因异质性（allelic heterogeneity）——在同一遗传基因座上，由不同的突变等位基因引起的相同或相似的表型。

地中海贫血（thalassemias）——由突变引起的血红蛋白数量异常的疾病，例如 α 珠蛋白缺乏引起 α 地中海贫血，β 珠蛋白缺乏引起 β 地中海贫血。

点突变（point mutation）——指一个核苷酸代替另一个。

电穿孔（electroporation）——在 DNA 存在时，应用短时高压电脉冲作用于细胞以使 DNA 进入该细胞。

凋亡（apoptosis）——程序化细胞死亡。

定位克隆（positional cloning）——在关于某个疾病生化基础的信息未知时，利用基因定位来克隆致病基因的方法。

动粒（kinetochore）——指着丝粒上纺锤丝附着的结构。

独立原则（independence principle）——两个或更多的独立事件联合发生的概率是他们各自发生概率的乘积。

端粒（telomere）——在染色体末端含有 TTAGGG 六聚体重复序列的特化 DNA－蛋白质结构。

多基因病（polygenic diseases）——由多个基因相互作用导致的疾病（或性状）。其中每个基因都起相对微效的作用。

多聚核糖体（polysome，polyribosome）——翻译过程中多个核糖体附着在 mRNA 上形成的结构。

多聚腺苷化作用（polyadenylation）——mRNA 3′端添加大约 200 个腺苷酸残基，显然与 mRNA向核外转运及稳定有关。

多态性（polymorphism）——在一个群体中，两种或更多由遗传决定变化的表型以一定频率发生，使其中最罕见的表型不能单靠重复突变来维持。实际上，如果罕见的等位基因发生频率至少为 0.01，而携带该等位基因的杂合子发生的频率大于 2%，则认为此遗传基因座是多态的。

多效性（pleiotropy）——指单个基因或基因对在几种器官系统和功能中具备多种作用。

多因素遗传（multifactorial inheritance）——由多个环境因素与多个基因相互作用产生的性状。

二倍体（diploid）——在绝大多数体细胞中，染色体的数目是其配子数目（单倍体数目）的二倍，人类二倍体染色体数目为 46。

发育异常（dysplasia）——指细胞异常地形成组织及其形态学后果。

翻译（translation）——由特异 mRNA 序列指导合成多肽的过程。

反转录病毒(retrovirus)——RNA 病毒,能进入宿主细胞反转录成为 DNA 拷贝,并随后稳定地插入其染色体 DNA。

反转录酶(reverse transcriptase)——在一 RNA 模板上催化合成 DNA 的酶。

非整倍体(aneuploid)——指单倍体非整倍数的任何染色体数目。通常非整倍体是指单条染色体的额外拷贝(三体性),或缺少单条染色体(单体性)。由减数分裂或有丝分裂过程中染色体不分离所致。

费城染色体(Philadelphia chromosome, Ph^1)——22 号染色体的结构异常,为 22 号染色体与 9 号染色体远端间的相互易位,发生在慢性粒细胞白血病患者部分骨髓细胞中。

分离(segregation)——在遗传学上,减数分裂时等位基因的分离。因为等位基因占据同源染色体上相同的基因座,它们进入不同的配子,即它们分离。

分子杂交(molecular hybridization)——单链 DNA 或 RNA 通过 Watson-Crick 碱基配对与其互补单链进行杂交的能力。

复合杂合子(compound heterozygote)——在某一特定基因座有两种不同突变等位基因的个体。

共有序列(consensus sequence)——诸如 CCAAT, TATA 或剪接供体这样的核苷酸序列,在基因表达中起重要的功能性作用。这些序列在不同基因间可能不尽相同,但它们趋于共有一定的特点,称为共有序列。

寡核苷酸(oligonucleotide)——指一段短的 DNA 或 RNA 片段,长达几个至一百个 bp,常规用专门设计的自动设备生产 DNA 寡核苷酸。

寡基因病(oligogenic diseases)——由相对较少的基因效应导致的疾病或性状,其中某些基因有相当大的作用。

关联(association)——在一个群体中两种或更多不同的表型同时发生的概率通常高于预期值(不同于连锁)。

合子(zygote)——由单倍体的男性和女性配子融合形成的双倍体细胞。

核酶(ribozyme)——能催化包括 DNA 裂解在内的特异反应的 RNA 分子。

核糖体(ribosome)——由核糖体 RNA 和蛋白质组成的细胞质细胞器,是 mRNA 赖以合成多肽的场所。

核小体(nucleosome)——是染色质的基本结构单位。由 146bp 的 DNA 包绕一 8 个组蛋白分子的核心组成。

核型(karyotype)——一个个体的染色体组成(染色体数目、性染色体组成和任何形态或数目的异常),也是标准划分一个个体的染色体显微照相及其制备过程(核型分析)中的常用术语。

候选基因(candidate gene)——一个基因定位于与某一特殊疾病相关的已知区域,其蛋白质产物的特征提示它可能是与该疾病相关的真实基因。实际上,若一疾病基因的物理位置已局限到含有限数目基因的狭小区域,那么定位于这个区域的所有基因都可称为候选基因。

互补(complementary)——由氢键结合的核苷酸嘌呤、嘧啶碱基对间的相互作用(A-T, G-C),用以形成核苷酸几何学的二级结构。

互补 DNA(complementary DNA,cDNA)——利用反转录酶,以 mRNA 为模板合成的 DNA。

互补分析(complementation analysis)——用以检测形成相似表型的两个突变是否为等位基因的遗传学试验。

环染色体(ring chromosome)——一种结构异常的染色体,其中每条臂的末端缺失,断臂重新连接成环。

回文结构(palindrome)——一段 DNA 序列的两条链含有相同的 5′至 3′的序列,例如 5′GAATTC3′3′CTTAAG5′。

肌营养不良蛋白(dystrophin)——由 X 连锁基因编码的蛋白质,其变异可引起杜氏及贝氏肌营养不良。通常认为这种蛋白质与横纹肌和心肌中收缩器固定于细胞膜有关。

基因间 DNA(intergenic DNA)——指转录的基因之间相当长的 DNA 片段,功能未知。

基因流(gene flow)——由于移民或通婚引起基因从一个群体向另一群体的逐渐扩散。

基因型(genotype)——指一个个体的基因组成,更特异而言是指某特定基因座的等位基因。

基因治疗(gene therapy)——通过向患者细胞中导入重组核酸序列来治疗人类疾病。

基因组(genome)——含有一个生物体全部遗传信息的完整的 DNA 序列。

基因组印刻(genomic imprinting)——是指依赖于遗传了父亲还是母亲的遗传信息差异表达。印刻被认为反映了在父母配子生成过程中差异形式的激活或失活。

基因座控制区(locus control region,LCR)——通常以组织特异性方式调节一簇基因表达的 DNA 节段。LCR 可与其调节的基因相距直到 10~100kb 远。

基因座异质性(locus heterogeneity)——指在不同基因座上的突变引起相似的表型。

畸形(malformation)——由内部异常发育过程导致一个或部分器官的形态缺陷。

畸形学(dysmorphology)——指临床遗传学中与先天性解剖异常的诊断和处理有关的领域。

剂量补偿(dosage compensation)——用以解释在男性和女性中 X 染色体上某一基因编码产物总量相同这一现象的机制。

加和原则(additivity principle)——如果两个事件相互排斥,那么获得其中一个或另一个的概率为它们的各自概率之和。

加帽位点(**cap site**)——转录的起始位点。

家族性(familial)——指受累个体亲属比一般群体更常见的任一性状,可能与遗传或环境的原因有关。

家族性高胆固醇血症(familial hypercholesterolemia)——是一种常染色体显性疾病,因缺乏低密度脂蛋白(LDL)受体导致血清胆固醇和 LDL-胆固醇升高。

甲胎蛋白(α-fetoprotein)——由胎儿产生、见于羊水的一种蛋白质。在正常孕期其含量增高直至妊娠第 34 周,但在某些正常情况(如多胎)或病理情况(如神经管缺陷,死胎)可见其异常增高。

假基因(pseudogenes)——具有可表达基因结构的 DNA 序列,推测其一度曾具有功能,但却在进化过程中获得了一个或多个突变以致不能产生蛋白质产物。

假阳性率(false-positive rate)——未受累个体中出现的全部阳性结果的百分率。

假阴性率(false-negative rate)——表现某性状或受累某疾病的个体中出现的全部阴性结果的百分率。

减数分裂(meiosis)——发生于有性繁殖生物体生殖细胞中的一种特殊的细胞分裂方式。双倍体细胞产生含单倍体染色体数配子的过程中,共发生两次减数分裂:减数分裂Ⅰ和Ⅱ。在减数分裂Ⅰ期发生染色体数目减半。

剪接(splicing)——在成熟 mRNA 形成时切除内含子。

简并密码(degeneracy of the code)——若干密码子编码相同的氨基酸。

简单序列重复(simple sequence repeat, SSR)——由多个重复的短序列元件组成的 DNA 节段,以 2bp 为单位的重复称为双核苷酸,3bp 的为三核苷酸等等。更长的重复尤其容易带有多个不同长度的多态等位基因。多态的 SSR 是人类基因组计划遗传标志的主要来源。

碱基对(base pair, bp)——在双链 DNA 中,嘌呤与嘧啶以氢键互补结合(腺嘌呤一定总是与胸腺嘧啶配对,鸟嘌呤一定总是与胞嘧啶配对)。将这种氢键连结的残基对命名为 1bp,是用来衡量一段 DNA 长度的单位。

建立者效应(founder effect)——在一个快速扩增的群体中,祖先的小人群中有一个或更多的建立者恰好是某一突变基因的携带者,那么这一突变基因的高频率出现称为建立者效应。

交换(crossing over)——在减数分裂前期Ⅰ,同源染色体断裂并相互再结合而导致染色体节段的交换。

酵母菌人工染色体(yeast artificial chromosome, YAC)——一种在酵母菌中繁殖的克隆载体,能够携带长达 1Mb 的 DNA 大插入片段。YAC 含有在酵母菌中自我复制所需的控制元件。

近端着丝粒染色体(acrocentric chromosome)——着丝粒靠近染色体端部的染色体。

近亲,近亲的(consanguinity, consanguineous)——指源于一共同祖先的遗传亲缘关系(真正的共同血缘)。

近亲系数(coefficient of relationship, r)——表示两个人继承共同祖先某一特定基因的概率,或是他们继承于同一祖先的全部相同基因的百分比。

经验再现风险(empirical recurrence risks)——利用实际观察到的家系中某一性状出现的频率,而非关于该性状确切遗传模式的知识计算出的同胞或后代这一性状的再现风险。

聚合酶链反应(polymerase chain reaction, PCR)——用以扩增短片段 DNA 的技术。这种方法依靠两侧寡核苷酸 DNA 引物和 DNA 聚合酶作用下引物延伸的重复循环实现。

均质染色区(homogeneously staining region, HSR)——细胞遗传学可检查的染色体上一段扩展区域,由一间隔节段的 DNA 扩增许多倍形成,称作一个扩增子。存在于染色体外的相似扩增序列称为双微体。

可变的表现度(variable expressivity)——指一种遗传性状可变的严重程度。具有多效性作用的相同突变基因的个体常显示可变的表现度,是环境的效应或是修饰突变基因表达的其他基因作用的结果。

可读框(reading frame)——以一系列三联体阅读一段核苷酸序列可能的方式之一。一个可读框不含终止密码,因此可能翻译成蛋白质。

克隆(clone)——由单一祖先细胞经有丝分裂形成的一组遗传学上相同的细胞。

肯定杂合子(obligate heterozygote)——指临床上不受累,但基于系谱分析必定携带突变基因的个体。(讨论见原文)

扩增(amplification)——一段 DNA 序列多个拷贝的产生。

莱昂化(lyonization)——在胚胎发育早期,女性体细胞两条 X 染色体中的一条随机失活,结果女性成为嵌合体,对于一条或另一条 X 染色体来说每个细胞都是功能半合性的。

厘摩(centiMorgan,cM)——遗传距离的单位。在一既定的减数分裂中,两个基因座如有1%重组机会即为相距 1cM。

连锁(linkage)——指两个或更多非等位基因的共遗传,因为它们的基因座在同一条染色体上紧密相邻,这样减数分裂后它们保持联接的概率大于非连锁基因 50%的预期概率。

连锁不平衡(linkage disequilibrium)——指特殊等位基因的优先联合。例如,一种疾病的突变等位基因与其邻近基因座上特异等位基因联合的频率要高于偶然发生的预测值。

连锁图(linkage map)——由连锁分析确定的显示某一特定物种遗传标记相对位置的染色体图。它不同于运用连锁分析、细胞遗传检测以及物理方法制作的物理图和基因图。

联会(synapsis)——在减数分裂中,减数分裂前期 I 的同源染色体配对。

链终止子突变(chain terminator mutations)——导致 mRNA 翻译时提前终止的无义或移码突变。

罗伯逊易位(Robertsonian translocation)——两个近端着丝粒染色体在着丝粒处融合,短臂和卫星丢失而成的易位。

裸鼠(nude mice)——免疫缺陷小鼠,可供来自小鼠或其他物种(如人类)的肿瘤细胞生长。

帽(cap)——一种修饰过的核苷酸——7-甲基鸟苷酸添加到延伸的 mRNA 链 5′端,是 mRNA 的正常加工、稳定和翻译所必需的。

密码子(codon)——DNA 或 RNA 分子中的三个连续碱基的三联体,决定一个氨基酸。

敏感性(sensitivity)——用于诊断试验,当疾病存在时测试结果呈阳性的频率。

内含子(intron)——基因的一个节段,起初被转录成为 RNA,之后通过剪接其两侧的外显子序列将它从初级转录本上移去,成熟 mRNA 中不含有内含子序列。

内切酶(endonuclease)——在一内部位点切割 DNA 的酶。限制性内切酶可在特定核苷酸序列切割 DNA。

破损(disruption)——指一个原本正常的发育过程由于外因阻断或干扰而致形态缺陷。

葡萄糖-6-磷酸脱氢酶(glucose-6-phosphate dehydrogenase,G6PD)——催化磷酸己糖旁路途径第一步的酶。

启动子(promoter)——位于基因 5′端的序列元件。它们固定转录起始的位点,控制 mRNA

的数量，有时也控制组织特异性。

千碱基(kilobase, kb)——DNA 序列上的一千个碱基对。

前突变(premutation)——三联体重复区中等大小的扩展(例如脆性 X 染色体综合征中 CGG 的重复)本身并不引起疾病，但向后代传递中不稳定，并可在减数分裂时增加其长度(例如脆性染色体综合征中母亲的情况)。

嵌合体(chimera)——指由一个以上合子衍生的细胞形成的个体。在制备敲除小鼠时，携带替代目的基因的胚胎干细胞与正常供体胚胎的细胞相混合。由这些胚胎生产的小鼠是衍生于胚胎干细胞与正常供体细胞的混合体，因此被称为嵌合体。

嵌合性(mosaicism)——指一个体具有两种或更多的有遗传差异的细胞系，这些细胞系来源于一单个合子，却因突变或不分离而不同。

群体遗传学(population genetics)——研究群体中基因分布以及逐代传递过程中维持与改变基因和基因型频率的因素的科学。

染色单体(chromatids)——在染色体复制后、但在分裂后期之前，加倍的染色体以着丝粒相连的两条平行相同的染色体链。

染色体畸变(chromosome aberration)——染色体数目或结构的异常。

染色体疾病(chromosomal disorder)——由于全部或部分染色体增多或丢失而形成的疾病(如 Down 综合征，21 三体性)。

染色质(chromatin)——构成染色体的核苷酸和蛋白质。

人类基因组计划(Human Genome Project)——将人类基因组全部 30 亿对碱基制图并测序的计划。

绒毛膜绒毛取样(chorionic villus sampling)——一种用于妊娠 8～10 周的产前诊断方法。是在超声指导下从绒毛膜的绒毛区(绒毛小叶)抽吸胎儿组织用于分析。

三体性(trisomy)——指增加一条染色体拷贝或由三个拷贝替代了正常的两个拷贝的状况。

生态遗传疾病(ecogenetic disorder)——由共同的环境因素与疾病特异的遗传易感性相互作用导致的疾病(如当 α_1 - 抗胰蛋白酶缺乏时吸烟可引起肺气肿)。

生殖系基因治疗(germline gene therapy)——将遗传替换的基因有意或无意地导入生殖细胞(如卵细胞、精子或它们的前体细胞)。生殖系基因治疗具有改变后代基因的潜能，与体细胞基因治疗形成对照。

生殖系嵌合体(germline mosaicism)——指两种或两种以上有遗传差异的生殖细胞系，它们衍生于同一合子，却因体细胞突变或不分离而形成差异，也称为性腺嵌合体。

适合度(fitness, f)——指生育力的大小和进而参与下一代基因库的情况。对于一个特定的基因型，其存活到生育年龄的后代数量与野生基因型数量相关。

数目可变的串联重复(variable number tandem repeat, VNTR)——由 DNA 节段多次重复而成的一类 DNA 多态性。有时也指小卫星。VNTR 在概念上类似 SSR，但它是更长，更复杂的重复单位。多态的 VNTR 标记对基因定位非常有用，但在很大程度上已被 SSR 所代替。

双合子的(dizygotic)——两个不同的卵子与两个不同的精子受精的产物，即不同的双胞

胎。

双杂合子(double heterozygote)——指在两个不同基因座上分别存在一个突变等位基因的个体。

索引病例(index case)——即先证者。

肽链指纹图谱(peptide fingerprint)——蛋白质或多肽部分水解所得的肽色谱模式,这项技术也可用于 DNA 或 RNA。

探针(probe)——在分子遗传学中,通过分子杂交检测其互补序列存在的标记 DNA 或 RNA 序列,是能够在许多 DNA 或 RNA 序列的复杂混合物中识别所需克隆的试剂。

特异性(specificity)——用于诊断试验,在无病时测试结果呈阴性的频率。

剔除(knockout)——这个术语是指用同源重组方法使特异的靶基因失活。用这一方法制备的小鼠叫做剔除小鼠,大量基因的剔除已被完成,可作为相应的人类基因缺陷的有效研究模型。

体细胞(somatic cell)——不涉及生殖系的任何器官的细胞。

体细胞基因治疗(somatic gene therapy)——将新的 DNA 物质插入受累个体一特定组织,用此方法插入的 DNA 不进入生殖系。

体细胞杂交(somatic cell hybrid)——将不同细胞融合而成的杂种细胞(通常来源于不同的物种)。在人类基因定位常使用人与啮齿动物细胞的杂种。

条件概率(conditional probability)——两个非独立事件联合发生的概率为第一个事件的概率与假设其发生后第二个事件发生概率的乘积。

同胞关系(sibship)——指家族中包括所有同胞(兄弟姐妹)的关系。

同线性,同线的基因(synteny, syntenic genes)——发生于同一染色体的特征。存在于一物种同一染色体上一起发生的基因群在另一物种中以相似的群组保存,这被称为表现出"同线性保守"。

同质异形的(sequence, dysmorphic)——指来源于单一早先异常或机械因素的多个异常的模式,如 Potter 同质异形性。

突变(mutation)——指基因组 DNA 序列中任何持久的、可遗传的改变。

突变率(mutation rate, μ)——表示每代每生殖细胞每基因座发生突变的频率。

外显率(penetrance)——指突变基因可见或缺乏表达的全或无现象。对于一种显性疾病,通过确定固有基因携带者(杂合子)的百分比进行定量分析;对于一突变基因,则通过表现其表型的个体进行分析。

外显子(exon)——基因的转录区,存在于成熟 mRNA 中,通常含有编码信息。

网织红细胞(reticulocyte)——成熟红细胞的前体。

微小缺失综合征(microdeletion syndrome)——由小到用显微镜无法观察的染色体缺失所引起的综合征。参见邻近基因综合征。

文库(library)——大量重组 DNA 克隆的集合,其中基因组的或 cDNA 片段被插入到特定载体上。

无义突变(nonsense mutation)——DNA 单碱基替换产生终止密码子。

物理图(physical map)——用限制性酶切制图和物理克隆等直接的物理学技术确定的基

因和遗传标记在染色体上的实际位置。与遗传图相比，后者用连锁信息确定它们的相对位置。

系谱(pedigree)——起源于“pied de grue”或鹤足；是一种家系史的图解，表明家系成员、他们与先证者的关系以及他们关于某种特殊遗传病的状况。

细胞滋养层(cytotrophoblast)——通过绒毛膜绒毛取样，从胎儿绒毛膜绒毛获得的快速分裂的细胞，用于核型和DNA分析。

细菌人工染色体(bacterial artificial chromosome, BAC)——是一种重组的质粒，用以克隆长达50~200kb的DNA大片段，能够在宿主细菌胞体内繁殖。

细菌噬菌体(bacteriophage)——一种细菌病毒。在分子生物学中这些噬菌体用作克隆时的载体。

先天性代谢缺陷(inborn error of metabolism)——一种遗传决定的生化异常，即因一种特殊酶的缺乏产生可能有病理后果的代谢阻断。这个概念由Archibald Garrod于1908年提出，并因此开辟了生化遗传学领域。

先证者(proband; propositus; proposita)——索引病例；一个受累者，通过他来发现并调查一个系谱。

显性(性状)[dominant(trait)]——指那些在杂合子(即带有一拷贝突变基因和一拷贝正常等位基因的个体)中表达的性状。

线粒体DNA(mitochondrial DNA)——指在具有自己独立DNA的细胞器——线粒体中环形染色体上的DNA。mtDNA在每个细胞中有许多个拷贝，呈母系遗传，并以快至基因组DNA 5~10倍的速度进化。

限制酶(restriction enzymes)——从细菌中纯化的内切酶，能在特定的核苷酸序列切断DNA双链。每个酶用其来源的生物命名，例如，*Eco*RⅠ来于大肠杆菌(*Escherichia Coli*)RY13，在GAATTC序列切割。

限制性片段长度多态性(restriction fragment length polymorphism, RFLP)——DNA序列的变异改变限制性片段的长度。这些可能是简单的点突变产生或破坏一个限制性位点，或是长度可变的区域(也称VNTR，数目可变的串联重复)。RFLP为连锁分析提供方便的标记。

限制性图(restriction map)——以碱基对表示DNA序列的图，用以指明限制性位点的位置。

相(phase)——此处“相”意指确定两个连锁基因座上变换的等位基因位于相同或不同的染色体。

相互易位(reciprocal translocation)——两个非同源染色体间的物质交换。

相邻基因综合征(contiguous gene syndrome)——一种通常散发的综合征，具有一致而复杂的表型，与涉及两个或更多相邻基因座的小片段缺失有关。亦见微小缺失综合征。

携带者(carrier)——指拥有一个突变等位基因的杂合性个体，通常只在纯合状态才引起疾病。

锌指蛋白(zinc finger proteins)——含有一个锌原子的指状结构的转录激活子蛋白。

信使RNA(messenger RNA, mRNA)——多肽赖以合成的模板，是基因中的遗传信息与其终产物——蛋白质特定氨基酸序列间的重要联系。

性染色体(sex chromosomes)——X 和 Y 染色体。

选型婚配(assortative mating)———一种非随机的婚配方式。意指在某一特殊亚群中,一个成员更可能与该亚群中其他成员婚配,称之为阳性,而更少可能婚配的则为阴性。

选择(selection)——基于生物适合度的差异,环境因素对特定表型及其基因型的作用。

血红蛋白 H(hemoglobin H) ——β 珠蛋白亚单位的四聚体。这种不溶形式的血红蛋白积累于血红蛋白 H 病中。

血红蛋白病(hemoglobinopathies)——由血红蛋白的质或量的异常引起的疾病。包括地中海贫血(量的异常),以及由突变影响血红蛋白功能导致的各种血红蛋白质性疾病,如镰状细胞贫血。

血小板来源的生长因子(platelet-derived growth factor,PDGF)——由血小板或其他细胞生产的蛋白质,能强烈刺激细胞生长、分裂并参与正常的损伤修复。PDGF 基因与 sis 原癌基因是相同的。

亚中着丝粒染色体(submetacentric chromosome)——着丝粒多少有些偏离染色体中央的染色体。

羊膜穿刺术(amniocentesis)———一种产前诊断的方法,通常在妊娠 4 至 6 月抽取羊膜囊内婴儿四周的羊水进行。

阳性预测值(positive predictive value,PPV)——指检验结果异常的个体发生或将发展成被检疾病的概率。

药理遗传学(pharmacogenetics)——研究对药物反应差异的遗传学基础的科学。

一致性(concordant)———对孪生同胞表现出相同的表型或性状。

移码突变(frameshift mutation)——指涉及非 3bp 整倍数的缺失或插入的突变,这种突变可以改变基因的可读框,此移码突变 3′端的所有编码区因此被读作凌乱语言,通常很快遇到终止密码子。

遗传基因座(genetic locus)——指染色体上的特定位置或部位。

遗传密码(genetic code)——特定 20 种不同氨基酸的碱基三联体。

遗传漂变(genetic drift)——指基因频率的随机波动,在小群体中最为明显。

遗传筛查(genetic screening)——在群体基础上检测以确认一小组具有受累某一特殊疾病或传递某一特定遗传异常的高风险个体。

遗传图(genetic map)——通过遗传连锁信息确定基因和遗传标记在染色体上的相对位置。与之对应的物理图则指利用直接的物理技术,如限制性制图和物理克隆来确定基因在染色体上的实际位置。

遗传异质性(genetic heterogeneity)——不同突变可引起相似表型。等位基因异质性是指同一基因座的不同突变,而基因座异质性则指突变发生在不同的基因座。

遗传致死(genetic lethal)———一种妨碍生育的遗传性疾病,它并不一定引起疾病或死亡,但一定妨碍生殖。

遗传咨询(genetic counseling)———一种旨在为受累某种遗传病或具有患病风险的个人或家庭提供关于他们自身状况的信息,以及为带有风险的夫妇提供信息、帮助他们做出合理生育决定的交流过程。

异固缩(heteropyknotic)——致密而暗染的失活染色质。

异染色质(heterochromatin)——经胰蛋白酶和 Giemsa 处理呈暗染的染色质,由重复的 DNA 组成。

异态性(heteromorphism)——指染色体的正常变异。

异序性(heteroplasmy)——在体细胞组织中两种或两种以上有遗传差异的线粒体群。

易位(translocation)——两个或更多的非同源染色体之间染色体物质的互换,可能为平衡或不平衡的。

隐性(性状)[recessive(trait)]——指那些只有在突变基因是纯合性的个体(即携带异常基因的双倍剂量)中才有临床表现的疾病。

荧光原位杂交(fluorescence in situ hybridization,FISH)——利用由荧光标记可检测的核苷酸探针来鉴定染色体或染色体的基因座。这一技术可用于研究间期细胞和停留于中期的细胞的染色体。

营养缺陷性突变型(auxotrophic mutants)——因缺乏催化某一特殊代谢步骤的能力而依赖于外源性营养物质的突变细胞。

有丝分裂(mitosis)——核分裂的过程,分为五步进行(前期、前中期、中期、后期、末期)。一个细胞分裂形成在遗传学上与其相同的两个子细胞。

诱变剂(mutagen)——引起 DNA 改变而增加突变率的物理或化学因子。

域(domain)——蛋白质氨基酸序列中的一个区域,它等效于某一特殊功能区或基因的相应节段。

原癌基因(proto-oncogene)——癌基因的正常细胞同源物。这些基因通常参与细胞生长的调控,并在整个进化中保存下来。

原位杂交(*in situ* hybridization)——基因定位最直接的方法之一,是克隆的 DNA 序列与分散在显微镜玻片上中期染色体之间的分子杂交。

杂合性丢失(loss of heterozygosity,LOH)——从肿瘤细胞的某个染色体区域丢失一个等位基因,而对于该个体正常细胞来说这一区域是杂合性的,可以用能区分两等位基因的多态标记来检测杂合性丢失。

杂合子,杂合性(heterozygote,heterozygous)——指一对同源染色体上特定的基因座有两个不同等位基因的个体。

杂交、复性(hybridization,reannealing)——在分子遗传学中,指一条 RNA 和一条 DNA 链或两条不同的 DNA 链互补配对。在体细胞遗传学中,指两个体细胞融合形成一个杂种细胞,含有来自两个亲本细胞的遗传信息。

载体(vector)——重组 DNA 或 RNA 物质的携带者。在克隆中,载体常是自我复制的 DNA,如质粒、噬菌体或酵母菌人工染色体(YAC),可向其中插入克隆的 DNA 片段。在基因治疗中,载体用来将重组 DNA 有效地导入到靶细胞中。

再现风险(recurrence risk)——指已发生某遗传病的家系中同辈或后代其他成员再发生该病的概率。

早现遗传(anticipation)——指一种遗传性疾病在较早年龄发病或在连续后代中严重程度增加。

增强子(enhancers)——以顺式作用增强邻近基因转录的 DNA 序列,可在3′或5′端的两个方向作用于基因,还可与该基因相隔相当的距离起作用。

整倍体(euploid)——正常单倍体配子染色体数目的整倍数,在正常体细胞中染色体数为2N,或称二倍体。

脂质体(liposome)——人工合成的类脂颗粒,用来将 DNA 高效导入靶细胞。

脂质体转染(lipofection)——利用类脂将 DNA 高效地导入靶细胞的方法。

植物红细胞凝集素(phytohemagglutinin)——从红豆中提取的植物血凝素,用以凝聚红细胞和刺激淋巴细胞分裂;可用于制备外周血核型。

质粒(plasmid)——染色体外可独立复制的环形 DNA,通常带有抗生素抗性基因并可在细菌中繁殖,在 DNA 重组技术中用做携带克隆 DNA 片段的载体。

致畸物(teratogen)——引起或增加先天性畸形发病率的物理或化学因子。

中间着丝粒染色体(metacentric chromosome)——指着丝粒位于染色体中部的染色体。

终止密码子(termination codon;stop codon)——UGA、UAA 或 UAG 三个密码子之一,可导致蛋白质合成的终止。(又称无义密码子)

肿瘤抑制基因(tumor suppressor gene)——一种正常的细胞基因,当纯合性缺失时可参与肿瘤的发生。肿瘤抑制基因一般以隐性方式起作用,需丢失两个拷贝才引起肿瘤发生,相对而言,癌基因则以显性方式发挥作用。

重组(recombination)——由连锁的基因座通过座位间交换(断裂和重接)形成新的组合。

重组 DNA 技术(recombinant DNA technology)——是将某一生物体一个基因或部分基因的 DNA 插入到另一生物体的基因组中的遗传分析技术。

重组率(recombination fraction,θ)——在连锁分析中,显示两基因座重组的减数分裂事件所占的比率。

重组染色体(recombinant chromosome)——在子代中染色体出现双亲都没有的基因型,这是减数分裂中交换的结果。

主要组织相容性复合物(major histocompatibility complex,MHC)——6号染色体短臂上的人类白细胞抗原(HLA)基因复合物,因其具有高度多态性,是连锁研究的一个理想标记。

转化反转录病毒(transforming retrovirus)——携带附加 DNA 序列(通常是癌基因)的反转录病毒,能使受染细胞转化出恶性表型。

转基因(transgenic)——实验产生的植物或动物,其中外源性 DNA 被人工导入并稳定地整合于生殖系。就小鼠而言,"转基因"一词包括标准原核显微注射产生的小鼠和用胚胎干细胞技术生产的剔除小鼠。

转录(transcription)——在细胞核中由 RNA 多聚酶催化、以双链 DNA 为模板合成单链 RNA 分子。

转染(transfection)——将一特殊改变的基因或基因节段转移到真核或原核细胞中。

转运 RNA(transfer RNA,tRNA)——与核糖体协作,沿信使 RNA 模板将氨基酸转移到位,以实现从 mRNA 序列翻译至蛋白质序列。

着丝粒(centromere)——染色体的中部狭窄部分,是染色单体赖以相连和动粒附着的一段

异染色质区域。

咨询者(consultand)——指寻求或参考遗传咨询的人。

姊妹染色单体(sister chromatids)——细胞周期的 S 期(DNA 合成)形成的两条相同的 DNA 链,以着丝粒相连结,直至有丝分裂或减数分裂Ⅱ后期才分开。

自身免疫疾病(autoimmune disorder)——一种以存在的抗体直接对抗个体自身抗原为特征的疾病。

综合征(syndrome)——被认为有病理遗传性关联的多个疾病的模式。

组蛋白(histones)——在染色体中与 DNA 结合的蛋白质,富含碱性氨基酸(如赖氨酸,精氨酸),并在真核生物进化中实际上始终保持不变。

***Alu* 重复序列**(*Alu* repetitive sequence)——位于基因间或内含子 DNA 中的中等重复序列,含有限制性内切酶 *Alu* Ⅰ 的识别位点,这些序列长约 300bp,并在人类基因组中重复出现约 500 000 次。

Barr 小体(Barr bodies)——见于女性体细胞的性染色质(来源于失活的 X 染色体)。

Bart 血红蛋白(hemoglobin Bart's)——γ 珠蛋白亚单位的四聚体。这是在胎儿水肿(α 珠蛋白基因表达完全缺乏)中累积的血红蛋白的主要形式。

Bayes 定理(Bayes theorem)——应用条件概率整合所有相关信息,以评估风险个体受累某遗传病或为该病携带者的似然性,这一定理称为 Bayes 定理。解释见正文。

B 细胞(B cells)——是以产生体液抗体来应答抗原刺激的小淋巴细胞。

CCAAT 盒(CCAAT box)——在许多(但并非全部)看似对有效定量转录至关重要的基因中,位于其转录起始位点上游 75 ~ 80bp 的一个区域。

DNA 印迹(Southern blot)——由 Edward Southern 设计的一种技术。将琼脂糖凝胶电泳分离的 DNA 片段转移到硝酸纤维素膜上,用放射性探针与之杂交以检测特异性 DAN 片段。

DNA,脱氧核糖核酸(DNA, deoxyribonucleic acid)——是生命的基本分子,其多聚体构成真核生物的基因。

DNA 反义链(antisense strand of DNA)——指 DNA 双链中的非编码链,它与 mRNA 互补,是 mRNA 合成的模板 。

DNA 甲基化(DNA methylation)——甲基基团共价结合于 DNA,在真核生物中最常涉及胞嘧啶的甲基化,并与基因的转录水平降低有关。

DNA 聚合酶(DNA polymerase)——负责 DNA 复制的酶,通过以 DNA 双螺旋中每条互补链为模板而完成新链的合成。

DNA 连接酶(DNA ligase)——催化两个 DNA 片段再连接的酶。

DNA 重排(DNA rearrangements)——指 DNA 片段的体细胞重组。在免疫系统的细胞中,可变区(V)、多变区(D)及连接区(J)重排以形成功能性抗体基因。

Duffy 血型基因座(duffy blood group locus)——由 Donahue 于 1968 年定位至 1 号染色体上的第一个常染色体基因。

Giemsa 带型(giemsa banding)——在胰蛋白酶温和处理后用 Giemsa 染色获得的每条人类染色体独特的特征性亮带或暗带。

Guthrie 试验(guthrie test)——用于筛查苯丙酮尿症的细菌抑制实验。

Hardy-Weinberg 平衡(hardy-Weinberg equilibrium)——是描述平衡群体中基因频率和基因型频率相互关系，以及确定已知某一性状频率的群体中等位基因频率和杂合携带者频率的定律。

LOD 值(LOD score)——用来检测一组连锁数据能否表明两基因座连锁与否的统计学方法。LOD 值是有利连锁优势比以 10 为底的对数。习惯上，LOD 值 +3(1000:1 优势比)可作为连锁的证据；LOD 值 -2(100:1 相反的优势比)提示不连锁。

Maxam-Gilbert 方法(Maxam-Gilbert method)——用化学降解来确定核苷酸精确序列的方法。

P ——染色体短臂的名称，源于法文 *petit*。

q——染色体长臂的名称。

RNA 印迹(Northern blot)——与 DNA 印迹类似的、通过杂交检测 RNA 的印迹技术，可检出与探针互补的 RNA 片段的大小及含量。

RNA，核糖核酸(RNA，ribonucleic acid)——核糖核酸的多聚体，含有核糖而不是脱氧核糖，是 DNA 转录的产物。

RNA 聚合酶(RNA polymerases)——依赖 DNA 模板的转录酶，用核苷三磷酸前体合成 RNA。

Sanger 法(Sanger method)——测定一克隆 DNA 片段精确核苷酸顺序的酶学方法。

TATA 盒(TATA box)——在许多(而非全部)与转录起始明显相关的基因中转录起始位点上游 25 ~ 30bp 的一段保守序列。

T 细胞(T cells)——在胸腺影响下承担细胞介导应答抗原的小淋巴细胞。

X 常染色体易位(X-autosome translocation)——X 染色体和一条常染色体之间的相互易位。

X 连锁疾病(X-linked disease)——X 染色体上一个突变基因编码的疾病。

X 染色体失活(X-chromosome inactivation)——见莱昂化。

α_1 - 抗胰蛋白酶(α_1 - antitrypsin)——是抑制弹性蛋白酶活性的一种丝氨酸蛋白酶抑制剂，该抑制剂的缺乏(如 α_1 - 抗胰蛋白酶不足)将导致严重的慢性肺和肝脏疾病。

问题和答案

第一章　问题

*1.讨论下列临床病例中遗传与环境的相对贡献：

a.苯丙酮尿症患儿的智力低下

b.澳大利亚殖民地—英国农民面颊的鳞状细胞癌

c.Tay-Sachs 病患儿的智力下降

d.一位 52 岁男性吸烟者患心肌梗塞

e.一位中年高血压患者经新血压疗法一天后出现无力晕厥

2.同一基因突变是如何导致若干不同的临床疾病或不引起任何疾病?

第一章　答案

*1.所有疾病都应认为是遗传和环境相互作用的结果(虽然有时很复杂)。

a.苯丙酮尿症是由于编码苯丙氨酸羟化酶基因的遗传性缺陷造成的,婴儿血中苯丙氨酸明显增高,导致智力低下。通过饮食控制苯丙氨酸的摄入可以避免这一严重后果。有相同遗传组成的两个个体(如一种特定酶的缺陷)可能有明显不同的临床结果,这取决于环境因素。

b.在太阳照射地区由于过多暴露于来自太阳的的紫外线,高加索人的皮肤癌发病率大大高于黑人,这表明遗传差异(本例中指肤色的不同)能导致相同环境下疾病易感性的差异。

c.在 Tay-Sachs 病中,已糖胺酶 A 的遗传缺陷是该病不可否认的因素,而与环境无关。

d. 吸烟是一个已知的危险因素,但是血脂的遗传成分和心脏病的家族史也是重要的因素。

e.药物代谢的遗传差异正不断被认识,当一个病人对治疗出现各种反应时要考虑药物代谢的遗传差异。该遗传差异是肝脏的细胞色素 P-450 代谢系统对某些药物代谢率明显不同,包括常用的抗高血压药物。

2.像第二章和第五章详细描述那样,基因含有大量的编码信息,并非一个基因所有的突变都会完全破坏该基因的功能。所以,不同的突变可能导致不同的结果。人类血红蛋白基因就是最好的例子,这已在本章论述。人类 β 血红蛋白基因的一个突变是造成镰状细胞贫血的原因。然而,人类 β 血红蛋白基因的不同突变可能产生异常氧亲和力的血红蛋白而引起不同的临床问题。许多人类 β 血红蛋白基因的单纯突变根

本不引起疾病，因为其蛋白产物仍完全有功能。然而，基因的“剔除”突变导致另一种称为β地中海贫血，这将在第六章叙述。

第二章 问题

1.绘制双链 DNA 中蛋氨酸密码子(ATG)的示意图。标出下列结构：

a.5′端

b.3′端

c.磷酸二酯骨架

d.2′碳

e.“碱基”

f.符合 Chargaff 碱基配对原则的化学键

2.在男性正常精子发育过程中，精原细胞既作为自我更新的干细胞，也可作为发育成初级精母细胞的细胞来源。初级精母细胞进入减数分裂，在减数分裂Ⅰ期后成为次级精母细胞，后者接着进入减数分裂Ⅱ期发育为精细胞。精细胞发育为成熟精子无需进一步的细胞分裂。请写出下表中的每种细胞和细胞周期阶段中，每个细胞中全部 DNA 碱基对数、染色体数和染色质数，并写出每个的“倍体性”(单倍体、二倍体)。

	DNA(bp)	染色体	染色单体	倍体数
G_1 期精原细胞				
G_2 期精原细胞				
初级精母细胞				
减数分裂Ⅰ期后的次级精母细胞				
减数分裂Ⅱ期后的精细胞				

3.在 DNA 复制(或 RNA 转录)过程中，用适当的核苷三磷酸向原 DNA(或 RNA)链添加碱基。在3′端还是5′端添加新碱基？是什么决定添加 A、G、C、T(RNA 为 U)哪种核苷酸？如果合成的3′脱氧核苷酸被加入生长链将发生什么？

4.我们每一个人都是从双亲、4 个祖父母和 8 个曾祖父母继承遗传信息的收益者，这是一条真理。孩子仅携带来自他或她 4 个祖父母中的两个的遗传信息的机率是多少？例如：该儿童没有携带来自他/她的祖母和外祖母的遗传信息而只含有他/她的祖父和外祖父的遗传信息的机率是多少？(假设无交换发生)，该儿童的性别是什么？

5.下面的氨基酸序列代表了蛋白质的一部分，每种氨基酸序列最可能代表那种突变？

正常：　Glu-Cys-Met-Phe-Trp-Asp

突变 A：　Glu-Cys-Ile-Phe-Trp-Asp

突变 B：　Glu-Val-Cys-Ser-Gly-Thr

突变 C：　Glu-Cys-Met-Phe

突变 D: Glu-Met-Tyr-Val-Leu-Gly

第二章 答案

1.

2.

	DNA(bp)	染色体	染色单体	倍体数
有丝分裂				
精原细胞 G_1	6×10^9	46	46	二倍体
精原细胞 G_2	12×10^9	46	92	二倍体
减数分裂				
初级精母细胞	12×10^9	46	92	二倍体
减数分裂Ⅰ期后的				
次级精母细胞	6×10^9	23	46	单倍体
减数分裂Ⅱ期后的精细胞	3×10^9	23	23	单倍体

3.核苷酸总是加到3′端。在DNA复制中,亲代DNA链作为模板按碱基互补配对原则指导合成新的DNA链(图2-4)。在RNA转录过程中,DNA双螺旋中的无义链作为模板按碱基配对原则合成新RNA链(图2-6)。对DNA和RNA而言,3′端脱氧核苷酸将阻止链的进一步延伸,因为3′位置缺乏羟基将阻止下一个核苷酸附着。

4.孩子父亲的23对染色体每对之一是由孩子的祖父遗传来的,按独立分离原则,祖父的

23条染色体全部分离到同一配子中并传给这个孩子的机率为$(1/2)^{23}$。对孩子母亲的父亲也是同样情况。因此，孩子只遗传他祖父和外祖父的染色体的机率为$(1/2)^{46}$或1.4×10^{-14}。因为孩子接受其祖父的Y染色体故他是男孩。

5. 突变A　错义突变[ATG(蛋氨酸)→ATT、ATG或ATT(全部为异亮氨酸)]

   突变B　移码突变[半胱氨酸密码(TGT)的第一个T缺失——注意产生的氨基酸顺序揭示哪个密码子用于正常序列]。

   突变C　无义突变[TGG(色氨酸)→TGA或TAG(终止密码)]。

   突变D　移码突变[第一个密码子中的第一、二、三个核苷酸后插入A(谷氨酸,GAA或GAG)]。

第三章　问题

1. X连锁显性和常染色体显性性状在家系中均为垂直遗传。你能找到哪4个特征帮助你在一组均显示相同垂直遗传性状的家系中区别X连锁显性和常染色体显性遗传?
2. 在下列名单中选出最贴切描述下面标有a→f情况的名词
   - i. 等位基因异质性
   - ii. 多效性
   - iii. 表现度不一致
   - iv. 早现遗传
   - v. 近亲结婚
   - vi. 基因座异质性

   a. 一个65岁的男性除白内障外没有显示强直性肌营养不良的征兆。但他的25岁的女儿具有典型的肌无力和消瘦,最近她生了一个女孩患有严重肌无力和发育迟缓。

   b. 多样色卟啉症,一种常染色体显性卟啉生物合成先天异常,能引起光敏性皮肤水疱,腹痛,末梢神经病和精神病发作。

   c. 一个男人患有严重的脊柱侧突,皮下长有数百个神经纤维瘤,他的姐姐患有损害容貌的血管丛状的神经纤维瘤,他30岁的儿子长有一些Lisch节(虹膜上生有的良性小结)和腋窝雀斑。

   d. 一种罕见类型的常染色体隐性生长激素缺陷,仅出现在瑞士阿尔卑斯山区的某些与世隔离的小村庄内。

   e. 鸟氨酸转氨甲酰酶(OTC)在尿素循环中是肝脏的一个重要的酶。一种无义突变和缺失都能引起*OTC*基因缺陷,是致死性新生儿高血氨的主要原因。

   f. 色素性视网膜炎,视网膜变性的一种形式,以常染色体和X连锁方式出现。
3. 正确/错误:

   a. 对于一种遗传致死性疾病,一定在他/她到达青春期之前引起病人死亡。

   b. 一个体从双亲之一遗传了导致疾病的突变基因,基因在妊娠和出生时都存在,因此所有的遗传病都是先天的。

   c. 无论何时,一种常染色体显性疾病出现于两个正常双亲所生出的儿童,那么这种疾

病是新突变所致。

d.如果他/她是家系中惟一的患者,那么该患者患有遗传性疾病是不可能的。

4.一个妇女患有囊性纤维化,她是这个家系中惟一患者。她的母亲、父亲、女儿、外孙、兄弟、侄儿的携带者风险是多少?

5.一妇女是一种X连锁疾病——血友病携带者。她生儿子的机会是多少?其概率用 P_1 表示。她生出受累儿童的机会是多少?其概率用 P_2 表示。生出受累男孩的机会是多少?其概率是 $P_1 \times P_2$ 吗?为什么?

*6.一个血液病学者已经收集了25个无亲缘关系的患有镰状细胞贫血的儿童以及他们同胞的资料,在这25个家庭的总共165个孩子当中,共有60个患镰状细胞贫血。由于镰状细胞贫血是一种常染色体隐性遗传病,该学者预计25%或165个儿童中有41人患病,令她迷惑的是在这些家庭中好像有更多的儿童受累,你能帮助她解释吗?

第三章 答案

1.

a.在X连锁显性遗传(XLD)中无男性→男性传递;而常染色体显性(AD)受累男性可将性状平均地传递给他的儿子,1/2受累。

b.在X连锁显性遗传(XLD)中,受累男性的所有女儿将遗传该性状;而在常染色体显性遗传(AD)中平均仅有50%遗传该性状。

c.当研究一些异常家系时,出现的受累女性将会是受累男性的二倍,因为女性有两条X染色体,因此有二倍的机会获得位于两条X染色体其中一条上的疾病等位基因;在常染色体显性(AD)疾病的一组大的家系中,受累女性与男性的比例是1:1。

d.XLD女性的病情要比仅含一条X染色体的XLD男性轻。

2.

a.早现遗传

b.多效性

c.表现度不一致

d.近亲结婚

e.等位基因异质性

f.基因座异质性

3.

a.错误。遗传致死性被定义为抑制生育这一事实,不一定反映寿命缩短。

b.错误。先天性指出生时表型是否表现。

c.错误。双亲之一可能携带有基因但是没有外显,或者可能曾经涉及到一个不同的生物学的父亲。

d.错误。这种情况是常染体中隐性遗传病的惯例而不是例外(图3-15),并且在其他遗传形式也可能发生。

4.

母亲:100%

父亲:100%

女儿:100%(假设她的丈夫不是携带者)

外孙:50%

她的(非受累)兄弟:67%(由于其双亲均是携带者,正常儿童有1/3的机会是正常纯合子,2/3的机会是携带者)

她的侄儿:33%

5. 她生儿子的概率是50%,生受累儿童的概率是25%。她生受累男孩的概率是25%,此值不是前二者的乘积,因为这不是两个独立事件。

*6. 她已经被确定偏差导入歧途。由于所有这些家庭最初被识别是因为至少有一人患镰状细胞贫血,这在双亲是携带者的所有家庭中不是随机样本,因为遗漏了那些父母二人是携带者而无受累儿童的家庭。有些方法可以纠正这种确定偏差。一个简单的方法是首先单纯地抽出最初用来确定家庭的所有患者。那么正确的估计是在总共165 - 25 = 140人中有60 - 25 = 35人受累,这与预期的25%一致。

第四章 问题

1. 有两个相同大小的部落位于相邻的岛屿上,但是他们彼此并不知道对方的存在。在部落1中,某种常染色体隐性遗传性白化病的发病率为1/10 000;而在部落2中,该病的发病率是部落1的100倍。火山活动突然使两岛之间产生了一个小岛样的桥梁,于是这两个部落很快发现对方并且开始自由婚配(假设不存在对这种白化病不利的选择作用)。回答下列问题:
   a. 在岛间桥梁出现之前,每个部落内正常等位基因和白化病致病基因的频率各为多少?
   b. 在两个部落刚刚合并成一个部落时,新生成的部落内正常等位基因和白化病致病基因的频率各为多少:这时,白化病的发病率为多少?
   c. 经过一代后,新生成的部落内基因频率和发病率各为多少?
2. 在群体研究中,检测5 000个个体的MN血型(一种常染色体共显性遗传性状)发现:MM基因型有2 400人,MN基因型有2 400人,NN基因型有200人。M基因与N基因的频率各为多少?该群体是否处于Hardy-Weinberg平衡?
3. 一位芬兰妇女是Meckel综合征(一种罕见的常染色体隐性疾病,在芬兰她居住地区的新生儿中发病率为1/9 000)患儿的母亲,她丧夫后打算再婚。假定她的新婚丈夫是下列几种情况时,该妇女怀孕后再生Meckel综合征患儿的风险是多少?
   a. 来自与她同一地理区域内的无关个体
   b. 她前夫的兄弟
   c. 她的一级表亲

*4. 视网膜环状萎缩是一种常染色体隐性视网膜疾病,多发于儿童晚期和青春期。假定该病在某一特定的人群中的发病率为1/324。有九对视力正常的夫妇从该群体中分

离出来,迁移到遥远的地方定居,从而形成一个新的遗传独立群体。在这个新的独立群体基因池中不带有视网膜环状萎缩致病基因的可能性是多少?在这个独立群体中,该基因频率与原来的群体相同的可能性是多少?该基因频率高于原来的群体的可能性是多少?

5.每一种遗传致死性常染色体显性疾病的发生必然是一种新的突变,为什么?

6.下表列出了单卵双生、异卵双生、普通同胞对三种不同疾病表现的一致性比率。说明遗传因素和环境因素在每种疾病中可能起到的作用。如果遗传因素在这些疾病中起到了主要的作用,你能推测出所涉及到的遗传基因座的数目吗?

疾病	一致性		
	单卵双生	异卵双生	普通同胞
	%	%	%
A	100	25	25
B	50	6	6
C	10	10	0.1

*7.某一特定的临床综合征可以以常染色体隐性方式遗传;也可以表现为由未知环境因素引起的完全非遗传性的表型模拟。如果这一综合征的表型模拟可引起80%的疾病,那么有该病患儿的夫妇再生患儿的风险是多少?

8.进行性强直性脊椎炎为一种慢性炎症性关节炎,在HLA-B基因座携带*B27*等位基因的个体的发病率比一般人群高90倍。这是否说明强直性脊椎炎与HLA-B作为紧密连锁呢?

9.用正确的术语解释下列某种遗传性疾病的HLA研究中的现象:

连锁

关联

连锁不平衡

a.位于HLA-B基因座的特定等位基因见于75%患有某种遗传性疾病的个体中和4%的正常对照。

b.在有多个同胞患病的家系中,受累个体通常继承亲代的某一相同等位基因,而非受累个体则继承另一个等位基因。

c.在已知与HLA连锁的疾病中,HLA-A和HLA-B基因座上的两个特定的等位基因在受累个体中出现的频率高于正常对照约15倍。

第四章 答案

1.

a.部落1中,白化病基因频率为$\sqrt{1/10\,000}=0.01$;部落2中该基因的频率为$\sqrt{1/100}=0.1$(假定在岛间桥梁出现之前两群体处于Hardy-Weinberg平衡)。

b.因为两部落大小相同,所以基因频率为二者均值,即 1/2(0.01 + 0.1) = 0.055;发病率也为二者均值,即 1/2(1/10 000 + 1/100) = 0.00505。

c.正常基因的频率 = 1 − 0.055 = 0.945;白化病基因频率 = 0.055;发病率 = $(0.055)^2$ = 0.003(假定在新部落中为随机婚配)。

2. 5000 名个体的 10 000 个等位基因如表所示:

	M 基因	N 基因
MM 个体	4800	0
MN 个体	2400	2400
NN 个体	0	400
总 计	7200	2800

因此,M 基因的频率为 7 200/10 000 = 0.72;N 基因的频率为 2 800/10 000 = 0.28。如果该群体处于 Hardy-Weinberg 平衡,则产生$(0.72)^2 \times 5\ 000 = 2\ 592$ 名 MM 型个体;2(0.72)(0.28)(5 000) = 2 016 名 MN 型个体;$(0.28)^2 \times 5\ 000 = 392$ 名 NN 型个体。然而这与实际观察值不相符合。说明该群体不处于遗传平衡。

3.可以确定该芬兰妇女是 Meckel 综合征致病基因的携带者,并且在该妇女所在的芬兰地区的一般群体中 $q = \sqrt{1/9\ 000} \approx 0.0105$。

a.与该妇女无亲属关系的新婚丈夫可能携带有 Meckel 综合征致病基因的概率为 2pq,即 2(0.9895)(0.0105) = 0.0209。即使父母双方都是该病的携带者,他们的后代将有 1/4 的孩子患病,即出生患儿的风险是 0.0209 × 1/4 = 0.0052 或 0.52%。

b.该妇女的前夫一定是 Meckel 综合征致病基因的携带者,并且这个致病基因一定来自于他的父母的一方。因此,他的兄弟也有 1/2 的可能携带有这一致病基因。在这种情况下,再生患儿的风险是 1/2 × 1 × 1/4 = 1/8。

c.通过绘制假设系谱,围绕目的个体的家庭成员推算携带该基因的个体可以很好地回答这些问题。可以推出该妇女的一级表亲携带有与该妇女相同的致病基因的可能性为 1/8,所以再生患儿的风险是 1/8 × 1 × 1/4 = 1/32。

*4.该群体中,视网膜环状萎缩致病基因的频率为 $\sqrt{1/324} = 1/18$。利用二项式分布进行计算。九对表型正常的夫妇共计 18 人或者说 36 个等位基因。在这 36 个基因中有 1/18的基因为视网膜环状萎缩致病基因,17/18 的基因为正常基因。对于每一个基因来说将有 17/18 的可能性为正常基因。因此,对于 36 个基因来说都为正常基因的可能性为$(17/18)^{36}$ = 13%。36 个基因中只有一个突变基因的可能性是 36 × (17/18)(1/18) = 27%(前面的数字 36 表示 36 个基因中的任何一个都有可能是突变基因)。则 36 个基因中有两个突变基因的可能性为[36! /(2! 34!)](17/18)34(1/18)2 = 28%(前面的数字为 630,表示 36 个基因中有 2 个基因存在突变的不同排列组合形式,如等位基因 1 和 2,等位基因 1 和 3,等位基因 1 和 4,……等位基因 2 和 3,等位基因 2 和 4,……)。

(这样计算的结果并不精确,因为我们并没有排除两个突变基因位于同一个个体的情况。前面我们提到了夫妇 18 人表型正常,计算时就应该修正前面的数字为 630 - 18 = 612。因为不存在两个突变基因位于同一个个体,形式有 18 种)。因此,在这个新的分离群中不带有致病基因的可能性为 13%;与原来的群体致病基因批女郎(2/36)相同的可能性为 28%;高于原来群体的可能性为 1 - (13% + 27% + 28%) = 32%。这说明在对小群体(遗传漂变)进行计算时,等位基因频率(甚至是有害基因)可以高于随机的预测值,这就是我们所说的建立者效应。

5.因为从定义上讲:遗传致死性常染色体显性疾病是不可以遗传的(否则它就不是遗传致死性的),所以一旦该病发生就一定是来自于新的突变。

(虽然这种情况通常是正确的,但也有例外:当亲代是携带有致病突变的性腺嵌和体时,他的后代可以由他继承下来遗传致死性常染色体显性疾病。在这种情况下,亲代可以生育并把突变传递给不止一个子女,而且患遗传致死性疾病。)

6.A 疾病中,单卵双生个体的遗传一致性极高,说明该病几乎完全由遗传因素决定,异卵双生和普通同胞具有相同的遗传一致性,约为单卵双生个体的 1/4,说明这种疾病在一个遗传基因座上以常染色体隐性遗传方式遗传。

B 疾病中,与异卵双生个体相比,单卵双生个体具有很高的遗传一致性,这说明遗传因素起主要作用。虽然在单卵双生个体间缺乏完全一致性,但提示环境因素也有重要作用。单卵双生与异卵双生一致性的比值约为 ~8,这提示有两个非连锁基因引起该病。

C 疾病中,单卵双生个体的遗传一致性很低,说明环境因素起主要作用。所有的双胞胎具有相同的一致性,而普通同胞的一致性显著低于异卵双生个体则说明遗传因素作用很小,而子宫中的环境因素是重要的影响因素。

*7.如果该病为常染色体隐性遗传的话,这对夫妇再生患儿的风险是 25%。如果是常染色体隐性遗传方式的概率只有 20%,则可以推出再生患儿的概率是 20% × 25% = 0.05。事实上,该家系的其他信息对更精确的推测十分有用。如,已知该夫妇已经生有 10 个正常的孩子,则该病是常染色体隐性遗传方式的可能性下降,而是表型模拟的可能性增加。

8.不能说明。只是 *B27* 等位基因与强直性脊椎炎相关。如果想要说明关节强直性脊椎炎致病基因与 *HLA-B* 基因座密切连锁,必须证实家系中 *B27* 基因座上等位基因存在共分离。在连锁研究中,与该病共分离的特定的基因只作为该基因座的遗传标记;不同家系中特定的等位基因也不同。与此相反,相关研究是依据所有或大部分受累个体同时具有已知作为上的特定的等位基因。

9.

a.关联

b.连锁

c.连锁不平衡

第五章 问题

1.编码一个酶的基因有 3 个外显子。下面显示的是前 2 个外显子有意义链的 DNA 序列(大写字母)及这些外显子周围的一些序列(小写字母)。

……accggcagtagATATCAGACCATGCTAATCGCTCCCCGACAGgtaagttgca……

atgaacgcaatatccttcctctcgacagGGGTAGTTT……

a.外显子 1 有多少碱基? 这个外显子编码该酶多少个氨基酸? 将它们写下来。(提示:在何处起始翻译。)

b.7-甲基鸟苷帽子能加在来源于该基因的 mRNA 的什么地方?

c.如果外显子 1 中第 14 个碱基由 C 变为 T, 该基因的表达将发生哪些变化?

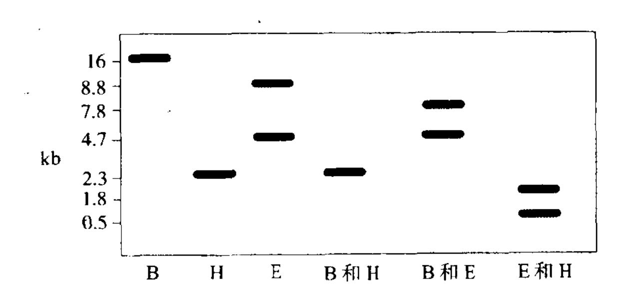
d.终止密码子在何处?

e.第 1 外显子中第 26 碱基由 C 转变为 T 将对该基因表达有何影响?

f.标记一个"D"和一个箭头,指向剪接供体位点; 标记一个"A"和一个箭头,指向剪接受体位点。

*2.在某个细胞提取物中你已经鉴定了某种酶活性,该酶能催化代谢产物 A 转变为产物 B。略述你能用到的三种不同策略来分离代表编码全部或部分该酶蛋白的 mRNA 的 cDNA。

3.以下显示的是用三种不同限制酶切割后,人 DNA DNA 印迹图。每个泳道用 B、E 或 H 标记,分别代表用限制酶 *Bam*HⅠ、*Eco*RⅠ和 *Hin*dⅢ切割。这三种酶或单独消化或其中两种共同消化。所用探针是人类基因组 DNA 的 2.3kb*Hin*dⅢ片段。画出该探针内及其两侧该限制酶两种可能的酶切位点图。



*4.

a.描述你所能用到的至少三种方法来有效地筛查一个大基因(100 kb,15 个外显子,全部 6 kb 编码序列)的点突变。

b.你将漏检哪些突变?

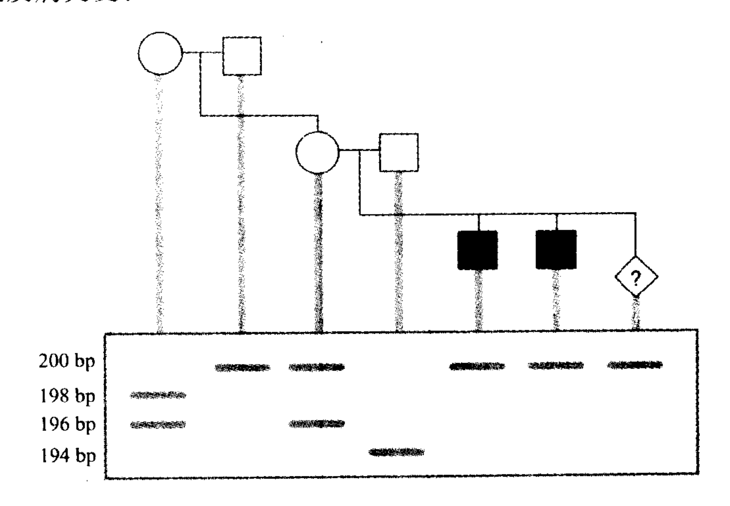
c.对于一个小基因(1kb,2 个外显子,500 bp 编码序列),你的方法有何不同?

d.如果在许多病人中检测出同样的点突变,你将怎样在众多的 DNA 样品中检测出该特定突变的存在与否?

5.你已经在一个甲型血友病患者的Ⅷ因子基因外显子 26 中检测出 C→T 的置换,

a.它就是导致该患者患甲型血友病的缺陷吗？除此之外，你还能解释该核苷酸的变化吗？

b.如果这个置换改变了一个氨基酸(一个 CGC 密码子变为 TGC，导致一个 Arg→Cys 的替换)，它能否改变你在问题 a 中的回答？你怎样来证明该突变是一个真正的甲型血友病突变？



6.上面系谱中实心方块代表患有严重甲型血友病的个体。空心方块代表非受累个体。每个个体的 DNA 都用 PCR 来扩增Ⅷ因子基因一个内含子的一个片段。系谱下面显示的是利用凝胶电泳分离 PCR 产物的结果。

a.什么原因导致凝胶中看到的不同大小的 DNA 片段？

b.在该家庭中你能找到突变来源于何处吗？

c.指出胎儿(菱形)的性别。

d.该胎儿患甲型血友病吗？

第五章 答案

1.

a.外显子 1 中有 31 个碱基，编码 7 个氨基酸：ATG 蛋氨酸；CTA 亮氨酸；ATC 异亮氨酸；GCT 丙氨酸；CCC 脯氨酸；CGA 精氨酸；CAG 谷氨酰胺。

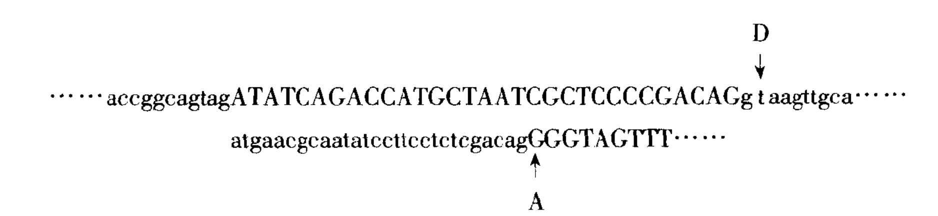
b.在原始 RNA 转录物 5′端第一个碱基处。从上面给的信息我们可以推断出 RNA 聚合酶在第一个大写字母 A 处起始转录。这就是将要加帽的地方。

c.TTA 仍然编码亮氨酸，所以这是一个静止碱基改变，对基因表达无作用。

d.TAG(外显子 2 中第 4～6 碱基)。

e.这个变化使编码精氨酸的密码子变为成熟前终止密码子，因此它在翻译时产生截短的蛋白。这对基因表达来说是非常有害的。

f.



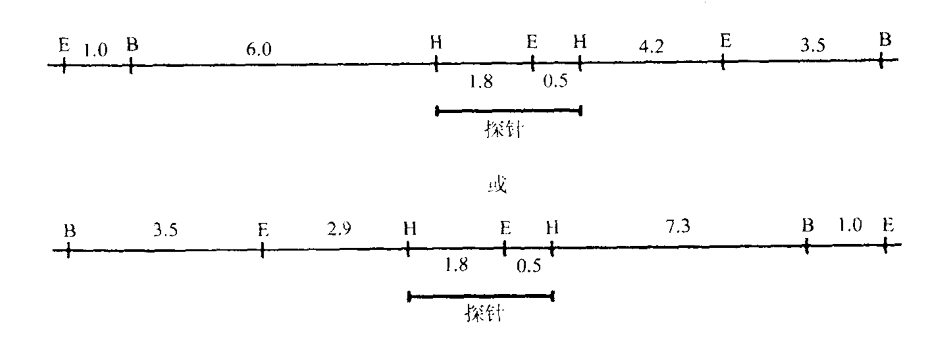
*2.

a.纯化该酶,得到氨基酸序列,推断编码这些氨基酸的核苷酸序列,最后用一组化学合成的、对应于这些序列的 DNA 片段作为探针来筛查一个 cDNA 文库。

b.纯化该酶,利用它作为抗原来制备针对该蛋白的特异性抗体。用此抗体来筛查利用表达载体所构建的 cDNA 文库。文库中携带克隆载体的细菌能够表达插入的 cDNA 序列。用特异抗体检测其相应的抗原来筛查该 cDNA 文库。

c.在一个系统中建立 cDNA 文库,该系统要求克隆的 cDNA 在培养细胞中表达。将文库中的克隆导入缺乏该酶活性的培养细胞后检测该酶活性的表达情况。

3.



*4.

a.筛查这么大的基因的全部序列通常是不现实的。因此我们应把精力集中在编码区和与其紧邻的剪接信号。许多含有单个外显子或外显子与内含子结合处的基因小片段能利用 PCR 的方法来扩增,进而用许多方法来筛查。这些方法包括变性梯度凝胶电泳(DGGE)、单链构象多态(SSCP)分析以及错配裂解方法等(见图 5-26)。另外一种可行的方法,如果表达目标基因的细胞容易得到,例如如果该基因在血细胞中表达,那么就可以利用 PCR 来扩增成熟的 mRNA,进而利用上述任何一种方法来分析或直接进行 DNA 测序分析。

b.化学错配裂解方法几乎能检测出 100% 的突变。尽管 DGGE 在一些系统中利用异源双链方法几乎能检测出所有突变,但相比之下 DGGE 效率要低得多。SSCP 在所有这些方法中可能最容易,但可能漏检 20% 或更多的突变。因为只筛查这个大基因的外显子和外显子与内含子交界处,故位于内含子内的突变不能检测出来。但这些突变有些能潜在地影响剪接或 RNA 加工的其他方面。位于启动子或其他转录调控信号内的突变也不能检测出来。

c.对于一个非常小的基因(大小与β珠蛋白相似)来说,对基因组DNA直接进行序列分析是最直接的方法,并且能几乎检测出所有的突变。

d.如果一个单个碱基改变导致了一个限制酶切识别位点的产生或丢失,就很容易利用PCR扩增结合限制酶消化和凝胶电泳分析来筛查。如果没有限制酶切位点改变,可以利用等位基因特异寡核苷酸杂交(ASO)来区分正常和突变的等位基因。需要注意的是ASO可以适用于所有的突变,包括那些改变限制酶切位点的突变。有很多其他快速和有效的筛查突变方法。与ASO一样,大多数这些方法都依靠敏感的杂交条件来区分即使是一个碱基的差别。

5.

a.尽管这很可能是导致甲型血友病的突变,但单个核苷酸的变化也可能是一个静止的DNA序列改变,并不导致氨基酸序列改变。

b.这个核苷酸改变导致了一个氨基酸改变。该事实进一步说明它很可能在该患者中存在缺陷。还值得注意的是这是一个非保守氨基酸的替换。同样要注意的是,这是在CpG二核苷酸处的一个C→T转换。而CpG是人类基因组突变的热点。然而这仍然可能是一个中性氨基酸变化,进而不影响Ⅷ因子的功能。

c.证明该核苷酸变化在这个患者中就是致病突变最直接的(也许是最明确的)方法是通过转染组织培养细胞以及检测重组Ⅷ因子的功能来表达突变蛋白。如果这个核苷酸变化导致了功能的丧失,这就进一步证明了该突变与该病人患病的一致性。遗传学证据也能说明问题。例如,在两个或更多的非相关血友病患者中检测出同一突变,而在一大组正常个体中不能检测出该突变。这就强有力说明该突变是致病突变。

6.

a.可以看到4种不同大小的谱带,被2bp长度单位区分开来。这是简单序列重复(SSR)多态性。它位于内含子内,由二核苷酸重复单位构成。

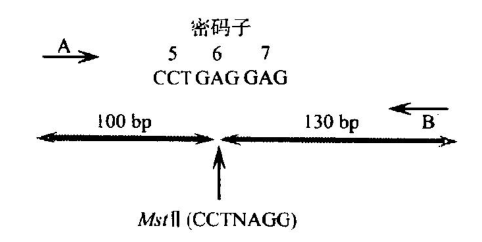
b.两个受累男孩每人从母亲遗传了一个200bp的等位基因。母亲是200/196杂合子。同理该母亲一定也分别从她父亲和母亲那儿遗传了200bp和196bp的等位基因。因为祖父没有患血友病,因此在两个患病男孩中与200bp等位基因相关的甲型血友病突变一定是一个新的突变(就像第七章所讨论的那样,大约1/3血友病患者是由于产生了新的突变)。这个突变最可能是在祖父辈减数分裂时发生的,该突变是约1/2严重甲型血友病患者常见的Ⅷ因子频繁倒位。

c.胎儿是一个男性。因为它只得到一个X染色体(200bp等位基因)。女性应当从父亲也遗传一个X(194bp等位基因)。

d.胎儿由于从母亲得到200bpⅧ因子等位基因,与两个受累儿子的等位基因相同,故胎儿也是甲型血友病的受累者。(如果母亲是一个生殖细胞嵌合体或在多态SSR标记与Ⅷ因子基因突变之间发生基因内重组,则该DNA诊断就会发生错误。但这种机率非常小。这些可能性在第七章和第九章中进行了讨论。)

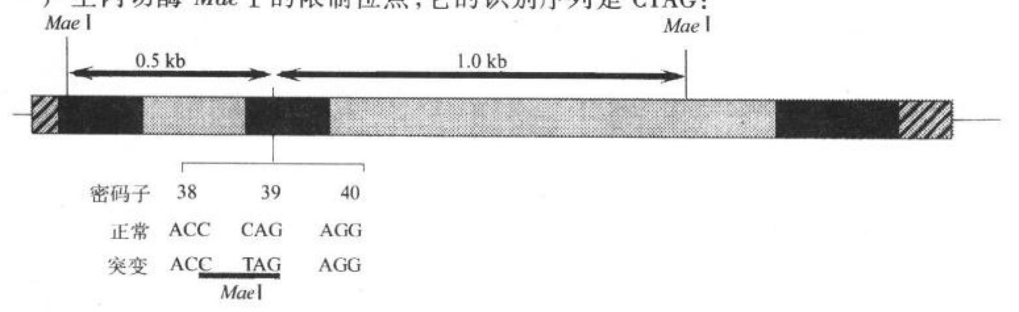
第六章 问题

1. 下列哪项策略对患严重型 β 地中海贫血的个体无益？
   a. 输血及铁螯合治疗
   b. 增加 α 珠蛋白产量
   c. 增加 β 珠蛋白产量
   d. 增加胎儿血红蛋白产量
   e. 降低 α 珠蛋白产量
2. 应用聚合酶链式反应方法(PCR,第五章)检测镰状细胞贫血。镰状突变改变了 β 珠蛋白基因第 6 密码子,由 GAG 变为 GTG,使正常情况下存在的 *Mst* Ⅱ 位点消失。用于 PCR 的引物如下所示：



   a. *Mst* Ⅱ 消化后,AA(纯合正常),AS(杂合"镰状性状")及 SS(纯合受累)个体的 PCR 产物其预期长度是多少？
   b. 血红蛋白 C(第 6 密码子为 AAG)纯合性个体及第 6 密码子移码突变的 β 地中海贫血(第 6 密码子的一个核苷酸缺失)纯合性个体的预期结果是什么？

*3. 你正在诊治一患有严重型 β 地中海贫血的病人。采集家族史时,你惊奇地发现,尽管母亲来自 β 地中海贫血很常见的希腊,父亲及所有他已知的祖先来自俄国。
   a. 一些可能的解释是什么？
   b. 对父亲进行亲权检查发现非常可能(> 99%)他实际上是孩子的生物学的父亲。然而他的血液显示绝无地贫性状(轻微型)的证据,这如何解释？
   c. 这对夫妇所生下一个孩子患严重型地贫、轻微型地贫的机会是多少？
   d. 基于这种情况,准备克隆病人的 β 珠蛋白基因。得到两个克隆,通过测序发现,一个在第 39 密码子处有一无义突变,是在地中海地区人群中常见的一种缺陷。突变产生内切酶 *Mae* Ⅰ 的限制位点,它的识别序列是 CTAG：



用 DNA 印迹法，如何确定这个突变的基因是否是从母亲处遗传下来的？具体说明你如何建立这个实验。假如你有母亲、父亲、受累的孩子的 DNA 及任何你想从 β 珠蛋白基因上克隆出的 DNA 片段，你如何用 PCR 方法进行检测？

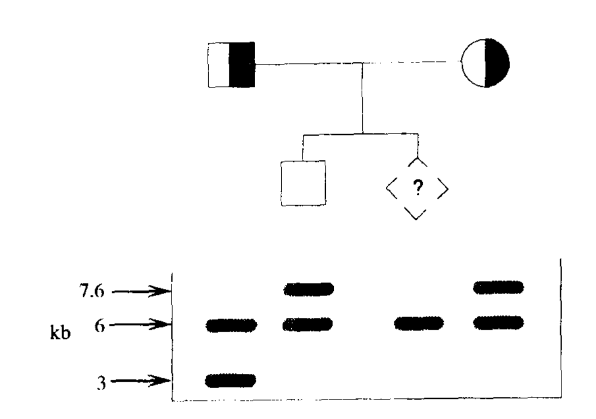
e. 另一个克隆没有这个同样的突变。假设它可能有一个以前未描述过的突变。你用头脑中丰富的洞察力，对此克隆测序。除内含子 2 末端有一核苷酸改变外全部 β 珠蛋白基因都正常：

	内含子 2	外显子 3
正常	—cctcttatcttcctcccacag	CTCCTGGC
突变	—cctcttatcttcctcccac [t] g	CTCCTGGG

这如何导致 β 地中海贫血？

你如何预测这是 β^0 地贫还是 β^+ 地贫等位基因？

4. 你正为一对希腊夫妇进行咨询，因为在人群普查中发现两人都是 β 地中海贫血携带者。他们的一个儿子检测结果正常（不是携带者）。他们想知道目前妊娠的胎儿是否受累。你通过羊水穿刺得到胎儿 DNA。用 β 基因 5′侧 10kb 的片段作探针，在此区域检测 *Hin*c Ⅱ RFLP。你得到下面的杂交结果：



胎儿受累了吗？

5. 一严重型 α 地中海贫血的受累胎儿发展成为严重的贫血及宫内并发症（胎儿水肿）。与此相反的是，严重型 β 地中海贫血的贫血症状一般直到出生后几个月才有明显的临床症状。为什么这两个相关的疾病的发病时间如此不同？

第六章 答案

1. B 增加 α 珠蛋白产量会进一步加重 α/β 不平衡。

2. 用引物 A 和 B，预期结果如下：

基因型	*Mst* Ⅱ 消化的 PCR 产物
AA	100,130
AS	100,130,230
SS	230
CC	100,130
△第 6 密码子/△第 6 密码子	230

*3.

a.可能的解释是:俄国的 β 地贫等位基因频率低,不正确的父子关系或新产生的突变。

b.这表明父源 β 珠蛋白基因在传递给孩子时获得一新的突变。

c.下一个孩子不太可能患严重型地贫,除非父亲是 β 地贫突变(见第三章)的生殖系嵌合体。因为母亲可能有轻微型地贫,下一个孩子患轻微型地贫的机会是 50%。

d.用 DNA 印迹法检测需要用 *Mae* Ⅰ 消化孩子和双亲的 DNA,电泳、转膜。然后用 β 珠蛋白基因序列如外显子 2 作探针进行杂交,第 39 密码子突变的杂合子个体,将有 0.5、1.0、1.5kb 带,而缺乏第 39 密码子突变的只有 1.5kb 带。可用 PCR 方法进行同样检测,设计的引物在密码子 39 两侧,扩增,*Mae* Ⅰ 消化,在凝胶上分析消化产物比较长度。

e.突变使内含子 2 末端正常的剪接受体消失,因为此位点的 AG(现突变为 TG)对正常剪接绝对必要,这可能会是 β^0 地中海贫血等位基因。

4.这类问题的关键是在每一双亲中确定标志 β 地中海贫血染色体的等位基因。正常的儿子必定是从每一双亲继承了正常的染色体,因此,父亲的 3kb 等位基因和母亲的 6kb 等位基因在他们各自的 β 地中海贫血染色体上。胎儿从父亲处继承一正常等位基因,从母亲处得到一 β 地中海贫血,因此,胎儿不受累(但将是一携带者)。

5.α 地中海贫血和 β 地中海贫血开始出现症状的相反时段可以用血红蛋白基因表达的发育模式不同(见图 6.2)来解释。从胚胎 ξ 基因到成年 α 珠蛋白基因表达的转换发生在胚胎早期,严重的 α 地中海贫血缺乏 α 珠蛋白基因导致无法形成血红蛋白 F(22)。相反,从胎儿 γ 到 β 珠蛋白基因表达的转换在生前开始,直到生后几个月才完成。只有到后期,当胎儿血红蛋白开始消失时 β 珠蛋白基因缺乏才产生临床症状。

第七章　问题

1.突变影响一种酶在红细胞和白细胞中表达的稳定性,为什么在红细胞中比在白细胞中更易引起可检测表型的变化?

2.假设一种酶是由一条常染色体编码的四条肽链组成的四聚体,你认为下列各个突变在一个杂合子个体中对这种酶的功能各会产生什么影响?

a.外显子 1 发生无义突变,导致翻译提前终止。

b.外显子1发生错义突变,使一个亚基虽可参与四聚体的构成,但会破坏酶活性位点。如果低于正常酶活性的10%足以引起疾病,你认为上述两种突变会产生何种遗传方式和疾病类型?

3.G6PD A^-与野生型基因相比有两个碱基发生了改变:376位核苷酸A→G, 202位核苷酸G→A,酶的迁移率改变,稳定性降低。你怎样用实验证实是哪种改变或两种改变共同导致迁移率改变和稳定性的降低?

4.用表7-1的资料计算 Pi^Z 和 Pi^S 的等位基因频率,人群中有多大比例为 Pi^{MS}?

5.哪两种疗法可有效地降低LDL受体缺陷杂合子的血胆固醇水平?

6.大约1/3的Ⅷ因子基因点突变引起甲型血友病,大约1/3的Ⅸ因子基因点突变引起乙型血友病,而引起G6PD缺乏症的错义突变大部分都涉及CpG双核苷酸中C残基的C→T改变。理论上讲,有多少双核苷酸可能是CpG? 为什么这些双核苷酸可能是人类基因组的突变"热点"?

第七章 答案

1.红细胞无核,依赖细胞分化过程转录和翻译产生蛋白质,不稳定的蛋白不能被替代。白细胞可以另外合成蛋白,所以总体酶活性可以保持行使正常功能的水平。而且,红细胞寿命较长(120天),而许多白细胞只能存活几天。

2.

a.该个体只能产生一半正常量的亚基。这些亚基可形成功能正常的酶。可以预见,他将会有正常酶水平一半的酶量。根据临床经验,这样的酶水平足以防止出现任何临床症状。

b.一半的亚基有这种突变,而产生有功能酶的惟一途径是酶的四个亚基全部正常,所以其发生机率应为 $(1/2)^4 = 0.06$。所以,他只有正常酶活力水平的6%。这样低的水平不足以维持正常功能。

无义突变引起的性状将以常染色体隐性方式遗传,而上述的错义突变可能引起常染色体显性遗传性状。

3.可用一个简单的步骤来检测这两个核苷酸改变引起的氨基酸变化。因为导致电荷差异的改变更易引起迁移率的改变,但稳定性较难检测。回答这个问题最可信的方法是克隆野生型和 A^- cDNA(见第五章),然后构建杂交分子,使分子中含一种或另一种 A^- 型突变,但不可两种都有,将这些杂种体在真核细胞中表达,这可以通过将cDNA导入"表达载体"然后将其转染任何一种组织培养细胞系来实现。这样就可对蛋白产物的稳定性和迁移率加以分析。

4.根据Hardy-Weinberg定律, Pi^Z 的频率是 $\sqrt{1/2\,500} = 0.02$, Pi^S 等位基因的频率是 $\sqrt{1/1\,000} = 0.032$,余下的是 M 等位基因的频率,即 $1 - 0.02 - 0.032 = 0.95$。Pi^{MS} 基因型个体的发生频率是 $2 \times 0.95 \times 0.032 = 0.06$。

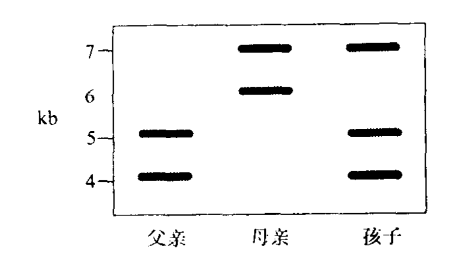
5.治疗高胆固醇血症最有效的方法是用称为抑制素的药物抑制内源性胆固醇的生物合成。抑制素是HMG-CoA还原酶的竞争性抑制剂。胆汁酸阻断剂对降低胆固醇水平

也有明显的效果。

6. 根据随机概率应只有1/16双核苷酸是CpG，实际上CpG双核酸在基因编码区的发生频率远低于这个值，也许因为这样的突变“热点”产生选择性的拮抗。但是，许多基因的5′上游区都可见CpG双核苷酸簇（CpG岛）。CpG双核苷酸被认为代表突变“热点”是因为胞嘧啶经常自发脱氨基形成胸腺嘧啶，引起C→T转换。

第八章　问题

1. 中央着丝粒染色体、亚中着丝粒染色体和近端着丝粒染色体之间有什么差别？近端着丝粒染色体上有什么特殊的结构和序列？
2. 下列哪些个体具有非整倍体的核型？
   a. 一个具有11q和22q平衡易位的女性
   b. 一个21三体型的Down综合征女患者
   c. 一个具有13和21号染色体的平衡罗伯逊易位的女性
   以上a、b、c三个个体各能产生几种配子？列出每一种配子与正常精子受精后产生的受精卵的染色体组成。
3. 为什么不是每一条人类染色体都在新生儿染色体三体中被发现？为什么孕9～10周的绒毛膜上皮细胞标本中染色体异常的类型和频率与它们在活产婴儿中的类型和频率不同？
4. 对一个已有一个21三体型Down综合征患儿的家庭进行了21号染色体的RFLP研究。这种RFLP有7、6、5和4kb这四种等位基因。下面所示的是孩子、她的父亲和母亲的DNA印迹杂交图。你能说出不分离发生在父母中哪一方？你能说出它发生在减数分裂的哪一阶段？



5. 你能预料下列个体有什么样的性腺和内外生殖器？
   a. 一个具有完全的*SRY*基因缺失的46,XY个体
   b. 一个46,XY的睾丸酮生化合成所必需的酶缺陷的纯合子
   c. 一个46,XY的睾丸酮受体缺陷的半合子
6. X连锁的凝血因子Ⅷ基因突变造成甲型血友病通常被人认为是仅有男性发病的一种疾病。然而，也存一些严重出血异常的甲型血友病女性。叙述这些妇女具有这种疾病的四种可能的机制。

7.如果一个个体是某一染色体非整倍体的嵌合体,你能得出哪些可能的推论导致嵌合性事件的发生?

8.对一个具有多发出生缺陷的儿童所做的常规染色体分析显示他具有一个非平衡染色体易位。为了判定易位涉及哪几条染色体,进行了 FISH 染色体原位杂交。对这种分析来说,下列哪一种 FISH 探针是最有用的:特异性基因座探针、着丝粒探针或全染色体探针?

9.为什么对一个临床明确诊断为 Down 综合征患者进行核型检查仍是必要的?

10.为什么对一个临床明确诊断为 Turner 综合征的患儿,获得其核型是必要的?是否对颊黏膜涂片进行 Barr 小体的计数就足够了?

*11.比较下列三联体动态重复所引起的疾病的发病机制:脆性 X 综合征、Huntington 病和强直性肌营养不良。

12.为 a~c 三种情况选择一个最能描述它们的术语。

Ⅰ 前突变

Ⅱ 基因组印迹

Ⅲ 单亲性二倍体

a.一个 7 岁的 Prader-Willi 综合征男孩,特征是中度智力低下,身材矮小,手足小和明显的巨食症。在分子遗传学分析中发现他有两个母源性 15 号染色体拷贝,而父源性染色体缺失。

b.一个 6 岁的 Angelman 综合征女孩,特征是严重的智力低下、癫痫和跳步,以及大下颌。在细胞遗传学分析中发现有母源性 15 号染色体的微小缺失。

c.一个临床表现正常的 32 岁妇女的 *FMR1* 基因 DNA 分析显示一个等位基因在第一外显子内有 21 个 CGG 重复单位,而另一个等位基因 CGG 重复次数为 92。她被告之生出患有脆性 X 综合征孩子的风险较高。

第八章 答案

1.不同之处是着丝粒的位置。在中央着丝粒染色体,着丝粒近似等分染色体。在亚中着丝粒染色体,两条臂大小不等。在近端着丝粒染色体,着丝粒接近末端。近端着丝粒染色体的短臂包含柄和随体,它是 rRNA 基因所在的位置。

2.b、c 是非整倍体。

a.有 6 种可能性:

具有 23 条正常染色体的配子;具有平衡的染色体组成的配子(衍生 11 号染色体和衍生 22 号染色体);含有一条正常 22 号染色体和一条衍生 11 号染色体的配子,与正常精子受精会产生 22 号染色体的部分三体型和 11 号染色体的部分单体型;含有一条正常 11 号染色体和一条衍生 22 号染色体的配子,受精后产生部分 11 三体型和部分 22 单体型;含有正常 11 号染色体和衍生 11 号染色体,受精后产生部分 11 三体型和部分 22 单体型;含有正常 22 号和衍生 22 号染色体,受精后产生部分 22 单体型;含有正常 22 号和衍生 22 号染色体,受精后产生部分 22 三体型和部分 11 单体型。

b.含有 23 条染色体(正常)或具有两条 21 号染色体的配子。与正常精子受精后,分别产生一个正常个体和一个 21 三体型的个体。

c.有六种可能:

正常配子;具有罗伯逊易位而无正常 13 号染色体或 21 号染色体的配子;受精后产生具有平衡罗伯逊易位的个体;具有(13;21)易位和正常 21 号染色体的配子,受精后产生 21 三体型;具有(13;21)易位和一条正常 13 号染色体的配子,受精后产生 13 三体型;仅有 13 号染色体而无 21 号染色体的配子,受精后产生 21 单体型;仅有 21 号染色体而无 13 号染色体的配子。

3.一些染色体异常与宫内发育不相容。在孕 9 ~ 10 周的 CVS 标本中能够见到某些染色体异常是因为这时由这些染色体异常引起的自发流产还没有发生。另一种可能性是某种异常的嵌合性可能产生携带异常细胞谱系的滋养层细胞。在这些细胞中,异常没有破坏胎盘功能,胎儿本身是正常的。

4.假定 RFLP 位点与 21 号染色体着丝粒紧密连锁。因为患儿从父亲处遗传了两个等位基因,而从母亲处仅遗传了一个等位基因,所以不分离是发生在父亲。而且,因为两条等位基因在父亲都出现,因此不分离事件发生在减数分裂Ⅰ期。

5.

a.这会产生性逆转。一个带有 Y 染色体的女性。

b.这会产生假两性畸形。性腺性别是男性(由正常 *SRY* 引起)而内外生殖器由于缺乏睾丸酮而不能发育成正常的男性结构。女性内生殖器也不会出现因为女性内生殖结构会发生退化,这是由男性性腺产生的非类固醇激素(Müller 管抑制物质)因子引起的。外生殖器通常是含糊的。

c.这会导致睾丸女性化。性腺性别是男性(由正常 *SRY* 基因引起)而内外生殖器由于缺乏正常的睾丸酮反应不能发育成正常的男性结构。女性内生殖器也不会出现,这是因为在睾丸分泌的非类固醇因子的作用下,女性内生殖器结构会发生退化。女性外生殖器正常。

6.

a.甲型血友病突变的纯合性。

b.一条 X 染色体突变的 Turner 综合征。

c.平衡的 X 常染色体易位,断裂位于 X 染色体上的血友病基因且正常 X 染色体失活。

d.合成Ⅷ因子的大多数细胞内正常 X 染色体的偶然失活,导致凝血因子水平低下。

最后,必须记住一种可能是另一种不同的疾病,它的临床表现近似甲型血友病,但以常染色体性状遗传。2N 型 von Willebrand 病,由 von Willebrand 因子的Ⅷ因子结合域突变所引起,能产生相似的临床表型,包括低血浆Ⅷ因子水平。

7.这一事件可能发生在妊娠后。

8.对判断哪些染色体片段交换,用来对整个染色体进行涂染的全染色体探针是必需的。着丝粒探针仅能提供有关一定染色体的着丝粒的信息,而特异性基因座探针只有当易位涉及含有这些基因座的染色体片段时才有用。

9.从遗传咨询的目的来看,了解 Down 综合征是 21 三体型还是易位是很重要的。如果是

易位造成的，父母和其他家族成员就可能是携带者。这样，再发风险就比父母核型正常而孩子是 21 三体型时要高得多。另外，也可能发现 21 三体型嵌合体，它也会影响预后判断和咨询。

10.计数 Barr 小体是不够的。有必要判定是否存在混合性腺发育不良中的含有 Y 染色体的细胞系嵌合体，因为具有 Turner 综合征表型和含有 Y 染色体细胞系的个体的退化性腺常有明显增高的癌症发病风险。通过精确的核型分析也可能获得一些判断预后的信息。

11.在脆性 X 综合征，外显子 1 的三联体动态重复导致 *FMR1* 基因的高甲基化和转录失活。相反，在舞蹈病，三联体重复涉及编码区，导致一个多聚谷氨酰胺的表达。这种改变了的蛋白可能与其他细胞成分相互作用而导致细胞异常和疾病。因此，这是获得功能突变的例子。在强直性肌营养不良中，三联体动态重复涉及 *DMPK* 基因 3′非翻译区。虽然这种突变可能改变 mRNA 稳定性从而影响该基因的表达水平，但确切的发病机制仍不知晓。

12.

a.单亲性二倍体。

b.基因组印刻。

c.前突变。

第九章 问题

1.检查下述体细胞嵌板，基因 *H* 出现在哪号染色体上？

在人与仓鼠杂种细胞中基因 *H* 与人类染色体的分离

细胞	基因	人类染色体																							
杂种	*H*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	Y	X
1	+	+	−	+	+	+	+	−	+	+	−	+	−	−	−	−	−	−	−	+	+	+	−	−	+
2	−	−	+	−	−	−	−	+	−	+	−	+	−	+	+	+	−	−	−	−	−	−	−	−	+
3	+	−	−	+	+	−	−	−	−	−	−	−	−	−	−	−	+	−	−	+	−	+	−	−	+
4	+	+	−	+	+	−	+	−	+	−	+	+	+	+	+	−	−	−	−	+	+	+	+	+	+
5	+	−	+	+	+	−	−	+	+	−	+	+	+	−	+	+	+	−	−	−	+	−	−	−	+
6	+	−	−	+	−	−	+	+	−	−	+	+	−	+	−	−	−	−	−	−	+	−	+	−	+
7	+	+	+	+	−	−	−	+	−	+	+	+	−	+	−	−	+	−	+	−	−	+	+	−	+
8	−	−	−	−	−	−	−	−	−	−	−	+	−	+	−	−	−	−	+	−	−	−	+	−	+
9	+	−	−	+	+	+	+	−	−	−	−	+	+	−	+	+	+	+	+	−	−	+	+	−	+
10	+	−	+	+	+	+	+	−	−	−	−	+	−	−	−	−	+	−	+	+	−	+	−	+	+

2.观察下列两个标记位点的 LOD 值表

将这些 LOD 值绘成图表。哪一个是 θ 的最大似然性估计值？该处的 LOD 值是多少？这两个标记是紧密连锁、松弛连锁、还是不连锁？你从这些数据中得到什么结论？

θ	LOD
0	$-\infty$
0.05	-10.32
0.10	-1.48
0.15	2.74
0.20	4.45
0.25	4.88
0.30	4.37
0.35	3.73
0.40	2.56
0.50	0

3.遗传性椭圆形红细胞增多症是一种常染色体显性疾病，该病的红细胞呈异常椭圆形且寿命短，导致在脾脏红细胞破坏增多而引起贫血。

在一个大家系中你可以观察到下述在 1 号染色体上 Rh 血型和遗传性椭圆形红细胞增多症间连锁研究的 LOD 值。

a.对于遗传性椭圆形红细胞增多症和 *Rh* 基因座的相对位置你能得到什么结论？

θ	LOD
0	$-\infty$
0.05	4.74
0.10	5.44
0.15	5.35
0.20	4.89
0.25	4.22
0.30	3.39
0.35	2.49
0.40	1.54
0.50	0

在第二个家系观察到的 LOD 值数据如下：

θ	LOD
0	-∞
0.05	-2.9
0.10	-1.44
0.15	-0.73
0.20	-0.34
0.25	-0.12
0.30	-0.01
0.35	0.03
0.40	0.02
0.50	0

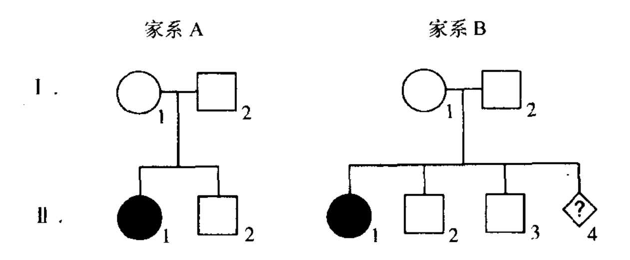
b.对于这个家系的遗传性椭圆形红细胞增多症与 Rh 血型基因座，你能得出什么结论？

c.从这两个家系所见的不同结果，给出两种解释？

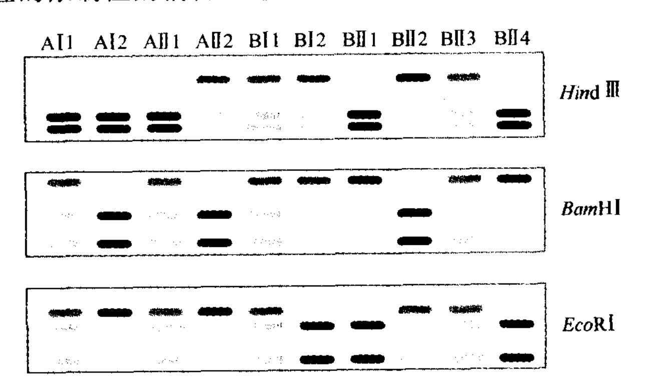
*4.假定 Hungtington 基因(*HD*)在 1983 年和 1993 年没被定位和克隆的话，请简单叙述用目前的方法和试剂你如何克隆 *HD* 基因。

5.至少列出 3 种人类疾病基因定位至特定染色体上的方法，而不用已知遗传标记的连锁分析方法。

6.A 和 B 是两个镰状细胞贫血家系。受累先证者(AⅡ-1 和 BⅡ-1)用黑色表示。



下面示 β 珠蛋白基因座内 3 个 RFLP 用 PCR 分析的基因型。跨越每个多态位点的 PCR 产物用相应的限制性酶消化后进行电泳。



a.判断每一个家系中所有个体的两个 β 珠蛋白基因的单体型。将特定等位基因的限制性位点的缺失标为“ - ”,未缺失标为“ + ”。

b.什么是镰状染色体的单体型？关于这些家系中镰形突变染色体的起源,你会得到哪些信息？

c.胎儿 BⅡ-4 是否患有镰状细胞贫血？你是如何确定的？你如何证实你的结论？

d.哪些个体是镰状细胞贫血携带者？你能确定这些吗？如何证实它们？

第九章　答案

1.基因 *H* 似乎定位于 3 号染色体,因为这是惟一具有完全一致性的染色体。

2.最大似然性估计值是 $\theta = 0.25$,此时,LOD 值为 4.88。由于 LOD 值大于 3,这两个标记明显连锁。但是,二者间也存在显著重组(25%)。

3.

a.在第一个家系中,疾病基因座与 Rh 连锁。最大似然性估计值 $\theta = 0.10$,此时,LOD 值为 5.44。

b.在第二个家系中,疾病基因座看来不与 Rh 连锁。

c.一个可能的解释是这个疾病有两个基因座,一个与 Rh 连锁,另一个不与 Rh 连锁(基因座异质性)。另一个可能性是数据太有限了。例如,在 10%重组处,不连锁的 LOD 值为 - 1.44,或者 27:1。然而,在大一些的 θ 值处,不连锁的机会不是那么高。这两组数据因相同基因座的突变代表了具有疾病家系的样本,但是连锁实际上不像第一个家系那么紧密。要注意 Rh 血型和遗传性红细胞增多症不是由相同的基因决定的。在$\theta = 0$处 LOD 值为 $-\infty$,表明在 Rh 血型基因和遗传性红细胞增多症基因座之间两个家系存在重组。

4.现在的方法寻找 *HD* 基因始于用一个跨越 10～20cM 的许多标记的嵌板进行基因组扫描。来自大 HD 家系中的 DNA,如图 9.26 所示,可以用 PCR 将这些标记进行基因型分析,再用专用计算机程序分析这些数据而将 *HD* 基因定位。该基因一旦被定位,再用已测定染色体片段上的其他标记将所有可能得到的 HD 家系分型,以便使候选间距缩至最小的可能区域。如果有 HD 家系且有大量标记的话,可能将此间距缩至 1cM 以内。所有已知位于此候选区域内的基因就可进行分析。必要的话,可用物理克隆方法获取跨越此候选区域的 YAC 和 BAC,还可以用各种技术识别位于这些克隆内的其他基因。通过 DNA 测序,比较患者和对照组 DNA 可以分析各种不同的候选基因。在所有患者样本都检出 *HD* 三联体重复扩展序列,而对照都无,这就有力证明已发现真正的 *HD* 基因。

5.

(1)X 连锁遗传。观察到一个疾病表现出遗传的 X 连锁类型,因此,将这个基因定位于 X 染色体上。

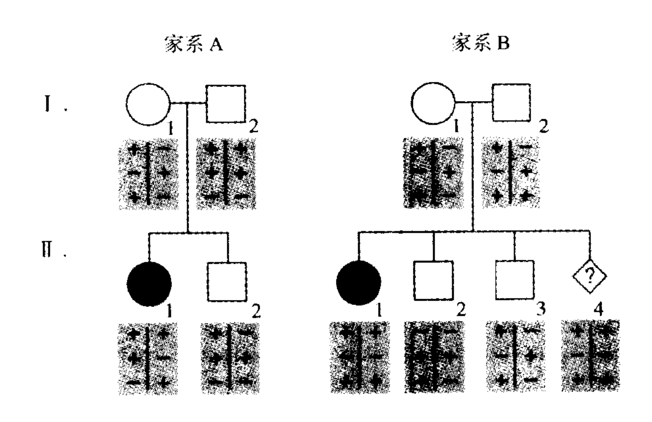
(2)通过鉴定某一患者在某一特定染色体上存在细胞遗传学方法可识别的缺失可以偶尔定位人类疾病。患有两种或两种以上遗传病的罕见患者,可以用多个紧密相

连基因的缺失来解释。(一组连续的基因综合征,正如第九章描述的患者 BB)。

(3)一个罕见患者出现一种特异染色体易位使致病基因断裂且引起表型改变也许可以将该基因定位。

(4)最后,对用功能克隆鉴定的许多人类疾病基因,用体细胞杂交或原位杂交技术如 FISH 分析,克隆序列可用于将基因定位于人染色体上。

6.



a.对于每一个限制性内切酶而言,多态限制性位点的消失("-")经限制性消化,电泳后,形成单一的大带,而该限制性位点存在时("+")这条大带将被消化而形成两条较小的带。杂合个体显示所有这三条带。纯合个体代表两个等位基因,他的泳带因含有双倍的 DNA 而显得更暗一些。因此,从电泳胶上我们可以直接推断出所有个体的基因型。对所有的酶而言,具有单一的、暗的、大的泳带的个体基因型为 -/-,只有两条弱带个体的基因型为 +/+,三条带都具有的个体基因型为 +/-。下一个问题是如何定单体型,也就是哪个 *Hin*dⅢ基因与 *Bam*HⅠ等位基因一同传递。例如,AⅠ-2是所有三个限制性酶的纯合子。因此,两个染色体必须具有单体型 + + -(*Hin*dⅢ、*Bam*HⅠ、*Eco*RⅠ)。对个体 AⅡ-1 而言,两个单体型的 *Hin*dⅢ为 +,而 *Bam*HⅠ为 + 和 -,*Eco*RⅠ为 + 和 -。表明单体型就是决定哪个 *Bam*HⅠ + 或 - 与哪个 *Eco*RⅠ一起传递。这可以通过观察双亲的单体型得出结论。AⅡ-1 必定继承他母亲(AⅠ-2)的 + + - 等位基因。所以,另一等位基因一定是来自父亲的 + + -。与此相似,当知道他父亲有一个 + - + 等位基因时,他的另一等位基因可以推断为 - + -。再看 AⅡ-2,两个单体型的 *Bam*HⅠ为 +,*Eco*RⅠ为 -,由于 *Hin*dⅢ是杂合性的,两个单体型必为 - + - 和 + + -。用这种方法,该家系的单体型就可确立。同样,BⅡ-1 是 + - + 的纯合子。由于每个双亲必须提供一个等位基因 + - +,你就可以推断 BⅠ-1′的另一等位基因为 - + -,BⅠ-2′的另一个等位基因为 - + +。你马上就会得知 BⅡ-2 为 - + + 和 - + -,BⅡ-4 和 BⅡ-1 一样是 + - + 的纯合子。BⅡ-3 有点难,因为他是所有三个酶的杂合子。然而,如果我们看一下他的双亲,*Hin*dⅢ一定与 - *Bam*HⅠ和 + *Eco*RⅠ连在一起,因此,他的两个单体型为 + - + 和 - + -。为了遗传检测或产前诊断,这就是正确解释在任何基因座

单体型是如何确定的。

b.简单看一下患者 AⅡ-1 和 BⅡ-1,两个镰状细胞染色体的单体型是显而易见的。BⅡ-1是 + - - + 的纯合子,AⅡ-1 在一条染色体上有相同的 + - - + 单体型,在另一条染色体上为 + + - 。这是一个连锁不平衡的例子。这种 + - - + 单体型看来是一种常见类型,四对父母中有三对与镰状细胞突变基因相关。这种 + + - 单体型可能代表突变与 *Bam*HⅠ、*Eco*RⅠ多态性间的一种重组形式。换言之,这可能代表镰状细胞突变的第二种独立发病形式,第一次事件发生于 + - - + 单体型背景,第二次发生于 + + - 单体型背景。事实上,似乎镰状突变确实有四种不同的单体型背景,这可能预示着四种独立的发育起源。

c.胎儿 BⅡ-4 与患儿 BⅡ-1 一样从双亲继承相同的两个镰状单体型染色体。因此,这个患儿极可能遗传了镰状细胞贫血。然而,为了绝对可靠,人们只需用 PCR-ASO 分析或限制性消化的方法直接查找突变的存在。

d.所有四个双亲是固有携带者(AⅠ-1、AⅠ-2、BⅠ-1、BⅠ-2)。此外,BⅡ-3 继承了母亲的镰状染色体,因此,可能也是一个携带者。BⅡ-2 继承了双亲正常等位基因而不是携带者。AⅡ-2 从母亲继承了一个 + + - 等位基因,但是,由于母亲的镰状染色体与她本身正常染色体是相同单体型,所以,人们不能区分这个孩子继承了哪一个。因此,这个个体有一半的机会是携带者。重申一次,真实突变可用上述方法评价。

第十章 问题

1.人类基因组计划主要争议之一是以一种组织方式解决人类基因组 30 亿碱基所预期的经济规模。例如,普通小型分子生物学实验室可以 5 美元的成本完成一对碱基的 DNA 测序。

a.按此效率,通过把小实验室的工作集合起来以“手工业作坊”方式进行人类基因组测序将花费多少?

b.在 1997 年,大多数效率高的基因组中心以大约 0.50 美元/bp 测序,那么现在人类基因组测序要花费多少?如果以此连续 9 年,每年的花费是多少?占美国 NIH 每年总预算(12 亿美元)的百分之几?

c.技术发展继续降低测序的成本,每个 bp 下降 0.01 美元,共节约多少钱?

2.定位克隆囊性纤维化基因(第十三章)需要几乎工作 10 年,由于人类基因组计划的完成,这类工作的进展将会更加快速。在下表中填写一个研究者在 1980 年(HGP 前)所必须进行的步骤和在获得完全序列后(HGP 后)还需要什么?

步　骤	HGP 前	HGP 后
收集家系		
连锁分析		
物理区域定位		
候选基因的鉴定		
突变的研究		

3.你想要在图 10-4 指示的"STS2"区开发一个 STS。

a.你将要设计什么引物(是 5'和 3'末端特异的)?

b.如果在人类基因组中是扩增一单个位点,那么你将如何检验那些引物?

4.你正试图构建含有一个重要疾病基因区域的物理图,你已做了 5 个 STS(S1 ~ S5),而你不知其次序,运用哪些 STS,你筛查一个 YAC 文库并发现 4 个 YAC 含有一个或更多的 STS。如下表所示:

YAC	S1	S2	S3	S4	S5
A	+	-	+	+	-
B	-	+	-	-	+
C	-	-	+	+	-
D	+	+	-	-	-

a.从这些资料构建一个物理图(按图 10-5 的版式)

b.你能说 YAC B 和 D、C 和 D、B 和 C 是重叠吗?

c.从这些资料,你能为这条染色体上的这个"重叠克隆群"定向吗?

d.如果 S4 和 S5 也是微卫星标记,而 S4 已表明在染色体 11q 上与 S5 相距 4cM,那么你能为此重叠克隆群定向吗?

5.在人类基因组计划刚开始时,某些争议是所有实验室应使用相同的 DNA 来源,理想的资源应该是非多态性的,即主要是单倍体。下述的来源是怎样的?

a.精子

b.Hydatiliform mole(一各无胎儿形态的异常妊娠,只是一个葡萄样的胎盘组织,通常由于妊娠时雌原核发生丢失引起,故核型为 23、Y,偶尔有转移发生)

*6.400 000 多人类 EST 的序列已被部分确定并储存于公共数据库中。分析这些序列表明它们来源于大约 42 000 个独立的转录物,这几乎是预期总数的一半。因为这些序列包含那些基因的部分或大部分的编码区,为什么我们需要一个人类基因组计划,对全部基因组 DNA 测序呢?为什么还要与这冒充物纠缠呢?

第十章 答案

1.

a.$ 150 亿

b.$ 1.5 亿;$16.7 亿,1.4%(注意:这是过高估价,因为美国能源部和其他国家也将测序)

c.$3000 万

2.

步　骤	HGP前	HGP后
收集家系	查明并仔细地确定系谱的表型	无变化
连锁分析	RFLP——慢、信息少、作图、2年的工作	完整遗传图并可容易地进行自动基因型分析
区域物理图	必须手工进行物理作图(一般为YAC),没有早先的界标,需1~2年	作图已毕并可公开使用
鉴定候选基因	为确定外显子必须筛查YAC或其他区域克隆、方法不完善,难以获得内含子/外显子结构,需1~2年	已经完成——计算机分析间隔序列,得到几乎所有的基因、内含子/外显子结构、有意义的同源序列
突变的研究	需要对正常人和患者的候选基因进行手工测序,若突变微小,经常需工作1~2年	由于自动化技术,突变检测极易进行

3.

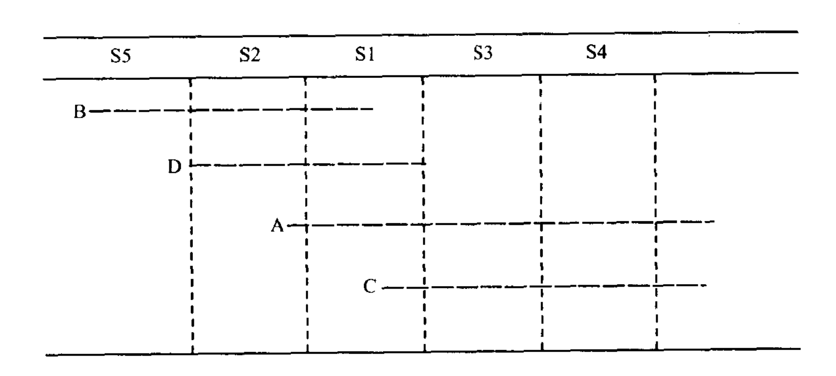
a.引物1:5′ GGATTCCGACTAGGTCGGCTT 3′

　引物2:5′ ATAGCCAATACCTAGCTAATCC 3′

b.为了解此引物是否工作,扩增人类基因组DNA(应用 *Taq* 聚合酶,见第五章),扩增产物电泳——你应该从该区的序列中得到预期大小的一个单带。为确认该序列不在其他染色体上(即重复的),还扩增一个只含有有意义染色体的体细胞杂种(第九章)和另一个含有一些其他人染色体但不含有STS2染色体的杂种细胞。前者出现预期大小的带,而后者应无产物(其实,偶尔的STS在这实验中仍有出现,并非惟一的带,因为它们来自单一染色体上的重复序列)

4.

a.



b.B和C肯定重叠

　C和D可能或不可能重叠

　B和C不能重叠

c.从上述资料看出,此叠连群没有方向性。

d.依此附加的信息,上图的顺序为“着丝粒向左,端粒向右”

*5.a.当一个精子是单倍体时,一个标本中的数百万精子足以制备基因组计划含有全部可能的等位基因的DNA——因此,由一次射精获得的DNA标本的多态性是不少于从同一男人的淋巴细胞获得的DNA标本。

b.一个23、Y hydatidiform 男人确是单倍体。然而,有理由相信也存在着体细胞突变,致使存活的组织具有这样一个异常的核型,所以这可能是预料出现在正常人基因组中一个危险的物质来源。

*6.有若干理由说明为什么完整的基因组序列是必需的,即使EST序列具有很大的实用性(特别在作图时)

a.cDNA含量低或只在难以得到的组织中表达,或者在发育过程中很少出现。

b.EST序列将不表现基因结构(内含子/外显子边界)或调节序列(启动子、增强子)。这个信息通常在理解一个基因在疾病中的作用是重要的。

c.对大转录子来说,其5'末端一般用EST策略是得不到的(因为此cDNA是在3'末端引发的,而且很难获得长于3~4kb的cDNA)。

d.EST序列将不表明有关染色体结构和进化,而且也不表明非转录序列(着丝粒、端粒、复制起始点等)的可能功能。

第十一章 问题

1.说出与原癌基因激活有关的4种机制。

2.解释为什么散发性视网膜母细胞瘤患者通常在单侧眼睛生长单个肿瘤,而遗传性视网膜母细胞瘤通常是发生在双眼或单侧眼睛多个部位发生肿瘤。

3.遗传肿瘤抑制基因的生殖系突变的患者发生癌症,从患者的正常和肿瘤DNA的分析证明肿瘤抑制基因内一个多态的杂合性丢失(LOH)。

a.列举3种可能解释这种LOH的机制。

b.从同一患者继发的独立癌中未能发现LOH,如何解释这一现象?

4.为什么DNA修复的遗传疾病似乎容易发生恶变?

*5.

a.一位25岁的妇女由于她有明显的癌症家族史向你咨询。她的母亲在35岁时因患乳腺癌做了乳腺切除术,并于50岁时死于卵巢癌。该患者的两个姨妈之一也于37岁时死于乳腺癌,列举对该患者进行DNA检测筛查*BRCA*基因突变可能的优缺点。

b.如果该患者是北欧犹太人,那会改变你的答案吗?

第十一章 答案

1.原癌基因激活所涉及的四种机制是:

a.点突变干扰癌基因产物的正常控制;

b.扩增导致癌基因产物的过表达;

c.染色体易位将一异源的启动子移到癌基因的上游,引起过表达。

d.染色体易位产生一杂种蛋白,它的一部分是癌基因产物,并缺少正常调节所需的蛋白质的一部分。

2.在散发性视网膜母细胞瘤中,同一细胞的两个等位基因必须都经过突变才发生疾病,因为不可能一个以上的细胞都经历这样罕见事件,所以肿瘤不可能发生于一个以上的病灶。相反,遗传型的源于单一突变,该突变影响已经有一个缺陷等位基因的细胞中惟一保留的正常等位基因。由于这第二次事件的概率足够高,因此在许多细胞能发生肿瘤的转化,并产生多灶性、双侧的疾病。

3.

a.在一个携带肿瘤抑制基因生殖系突变的患者中,正常等位基因失活的各种机制如图 11.5 所示。导致 LOH 的 3 种机制是:(i)正常染色体的丢失,(ii)染色体丢失和复制,和(iii)在正常染色体和携带突变基因的染色体之间的有丝分裂交换。

b.正常等位基因也可因独立的突变事件而失活。与另一个等位基因的生殖系突变相结合,尽管未见到 LOH,但结果是肿瘤抑制基因功能的完全丧失和肿瘤的发生。

4.一个合理的解释是,在这些疾病中见到突变率的增加导致细胞癌基因突变,使癌基因激活并引起肿瘤的转化。

5.

a.患者的母亲具有乳腺癌和卵巢癌的家族史,同时母亲和姨母在年轻时发生这些癌症,这有力地说明 *BRCA1* 和 *BRCA2* 基因突变。DNA 检测的优点包括更精确地评价患者发生乳腺癌或卵巢癌的风险。在患者中鉴定生殖系突变将提示在她一生中发生乳腺癌的风险为 60% ~ 90%,患卵巢癌的风险为 20% ~ 60%。如果能排除突变的存在,这将使她的风险降至总人口的风险水平(乳腺癌 11%,卵巢癌 1.4%)。虽然在高风险患者中尚无特异的干预措施有利于预防乳腺癌或卵巢癌,但有一些选择性手段包括经常性的监测和预防性手术。患者可能也希望得到这些信息以便安排生活计划。

缺点包括检测结果对情绪的潜在消极影响,昂贵的检测费用,对健康保险的可行性或费用以及就业所带来潜在的负面效应。注意到受累的亲属都无法进行检测这一点是很重要的。因此,在患者中未查到突变仅表明她的风险较低。阴性的检测结果可能意味着或是患者没有遗传的突变,或是该家族中的突变用现在的分析方法没有检测到。相比之下,如果在一个受累亲属中首先鉴定了一个突变,那么患者不存在此突变,将使她的风险降至一般群体的标准。正因为如此,宁可对家系中已知可能受累的成员进行检测分析、有时可以对次级亲属的病理标本进行 DNA 诊断性检测,虽然这方面研究仍主要限于实验室的研究。

b.正如第十一章所指出的,一个 *BRCA1* 和两个 *BRCA2* 基因的特异突变似乎可占北欧犹太人群中 *BRCA* 突变的大多数。因此,即使受累个体没有进行最初的检测,在此人群中阴性的筛查结果可能具有更大的价值。

第十二章 问题

1.与 Thomas Smith 的谈话:我和妻子 Alice 想要一个孩子,好像在我的家族中存在一种神

经疾病,我们想知道我们的孩子是否有患此病的任何风险?

我的哥哥 Don 在他少年时开始出现肌无力,脚部和腿部出现感觉丧失,十年前当他二十五岁时情况越来越糟直至他残废。当他的这些症状刚出现时,我母亲告诉我们这些症状看起来与我父亲的哥哥 Fred Smith(68 岁)和妹妹 Sally (63 岁)的情况非常相似,虽然我父亲 Tom 和他同父异母的兄妹——双胞胎 Eddie 和 Jane(73 岁)从未有此症状。

我的堂兄弟 Frank 和 Bill 看起来健康,他们已经二十多岁。他们的父亲——我的叔父 Fred Smith 是患有腿部肌无力和麻木的亲属之一。年龄稍大些的 Frank 已婚没有孩子,虽然我认为他妻子去年有过一次流产。我的姑母 Sally 有一个女儿 Linda,她看起来也有腿部肌无力问题,正担心她的两个孩子 Dick 和 Bob 是否也有该疾病。

我的父亲 Thomas Smith 在他 40 岁时死于一场车祸,他是家族中最年轻的。他的父母已故。他的父亲 William 健康,他在 80 多岁时死于某种心脏疾患。他的第一个妻子在生 Eddie 和 Jane 时死亡。他的第二个妻子 Marry,是我的祖母,我不熟悉她,但每个人都说她常抱怨脚有"针扎样"疼痛,最后她变得如此衰弱以致坐在轮椅上。

Alice 的家族非常健康,她的父母仍健在,她的三个妹妹和二个兄弟也都健在。

画出上述家族的系谱。依据上述特点,你想一想上述特征是何种遗传方式? 你从获得家族史的疾病中还学到了什么?

2.为什么仔细的体格检查和询问家族史是妥善对待伴先天缺陷新生儿的重要部分?

3.从下列表中选择最合适的诊断和试验方法,使之适用于从 a 到 g 的各种疾病的产前诊断。数字标记的词组可用一次以上,一种以上数字标记的词组可以用于每一种用字母标记的疾病。

1) 羊膜穿刺术
2) 绒毛采样
3) 超声诊断
4) 细胞遗传学分析
5) DNA 印迹
6) PCR
7) 等位基因特异寡核苷酸杂交
8) 利用基因内多态进行的连锁分析

a.先天性强直性肌营养不良
b.X 染色体脆性综合征
c.脊髓脊膜突出
d.甲型血友病
e.Huntington 病
f.β 地中海贫血
g.13 三体型

4.Jane 是一名 26 岁妇女,她的哥哥死于囊性纤维化。她的丈夫 Dick 也是北欧血统,无亲缘关系。假设囊性纤维化(CF)在北欧人群中的发病率为 1/2 500,Jane 是囊性纤维

化携带者的概率有多大？在任何一次妊娠中，他们的子女患囊性纤维化的概率有多大？

Dick 和 Jane 决定做携带者筛查试验，检测 70 个较常见的突变，可鉴别北欧人中 90%的囊性纤维化携带者。假设该试验无错误结果，Jane 试验结果为 70 个 CF 突变均为阴性，她是携带者的风险有多大？Dick 的试验结果为阳性，那么 Dick 和 Jane 任何一个子女的患病风险有多大？

Jane 和 Dick 离异分居，Dick 与 Jane 的一级堂姐妹 Laura 结婚。Laura 是携带者的风险有多大？Dick 与 Laura 的任何一个子女患囊性纤维化的风险有多大？

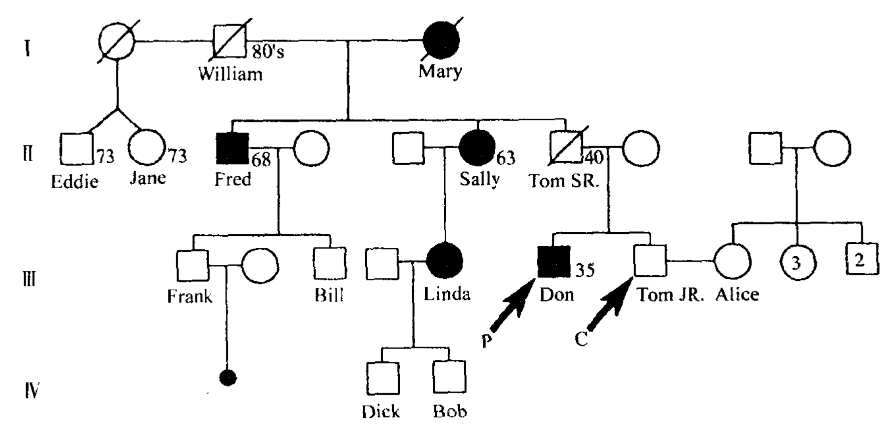
5. 在 Ann Arbor(大约 100 000 人)人群中，进行随机筛查试验以检测一种严重的常染色体隐性遗传病携带者。这种疾病在一般人群中的发病率大约为 1/10 000，该试验在大商场内进行。

一名 24 岁男子向你咨询，他感到极度心烦意乱。他已参加了该试验并被告知试验结果阳性。尽管没有家族史，但他已确信自己是一名携带者。当你试着安慰他时，他说："不用麻烦了，我已被告知这个试验有 98%敏感性和 90%的特异性。有那样水平的敏感性，这个结果一定是正确的！"这一男子不是所检测疾病的携带者的概率有多大？

一名 20 岁女子，她的一个兄弟患有此病，已作了同样的筛查试验，并向你咨询。她是携带者试验中 300 个未患病而他们的同胞患有此病的个体之一。她的携带者试验结果为阴性，她想知道这是否说明她真的不是携带者。她实际是携带者的概率有多大？

第十二章 答案

1. 在这个家系中描述的神经症状是以常染色体显性方式遗传的疾病。在发病的严重程度、发病年龄(可变表现度)方面表现相当不同。Tom 先生 40 岁死于车祸，未表现出此病任何体征的事实，说明他的发病年龄比他的儿子 Don 要晚。尽管后代中发病年龄



提前,但人们并不能得出根据这个单一家系是否可描述预测的结论。标有 P 的箭头代表先证者 Don,第一个使其家族受到医学关注的个体,标有 C 的箭头是咨询者,一个寻求遗传咨询的个体,虽然此时他还尚未表现出受累此病。黑色标记代表受累此病的个体。

2.很显然,仔细的体格检查对评价出生缺陷的自然状况、严重程度、对尝试做诊断、进行适当的手术和/或其他治疗是必要的。然而,同样重要的是检查是进行准确诊断的重要部分,对如下情况提供证据是重要的:是单一性先天畸形,还是多发性先天畸形;多发性是否是呈因果关系或致病关系,是顺序发生还是综合征。一个准确的诊断对提供遗传咨询,为双亲提供是否在以后的妊娠中会有此病再发风险的信息,为也有此风险的其他亲属提供遗传咨询和干预是非常重要的。体格检查与仔细询问家族史相结合也可帮助决定哪一种诊断研究,例如细胞遗传学或分子分析法,对患者和其他家族成员的风险作出诊断和评价是重要的。

3.a.1、2 和 5。

b.1、2 和 5。6(PCR)对描述前突变状态的特征有用。

c.1 和 3。

d.1、2、5 和 8。

e.1、2 和 6。

f.1、2 和 7。

g.1、2 和 4。

4.患有常染色体隐性遗传病,例如囊性纤维化的患者未受累同胞是携带者的概率为 2/3。Dick 无囊性纤维化家族史,根据囊性纤维化人群发病率判定他也有患病风险。囊性纤维化的发病率为 1/2 500,基因频率(将所有的 *CF* 突变一起考虑)为 $\sqrt{1/2\ 500} = 0.02$ 或1/50。假设 Hardy-Weinberg 平衡。携带者的频率为 $2pq = 2 \times (0.02) \times (0.98) = 0.04$ 或1/25。Jane 和 Dick 都是携带者的概率为$(2/3) \times (1/25)$,在每一次妊娠中受累胎儿的风险为 $1/4 \times (2/3) \times (1/25) = 1/150$ 或接近于 0.067。

考虑到 Jane 和 Dick DNA 检测时可能出现的另外情况,需利用 Bayes 定理。条件概率的计算在以下列出。

	Jane 是携带者	Jane 不是携带者
前概率	2/3	1/3
条件概率(阴性试验)	1/10	1
联合概率	2/30	1/3 或 10/30
后概率	$\frac{2/30}{2/30 + 10/30} = 1/6$	$\frac{10/30}{2/30 + 10/30} = 5/6$

假设 Jane 试验结果阴性,她是携带者的概率为 1/6 或 0.17。Dick 试验结果阳性,是携带者。因此任何一个子女的风险为 $1/4(1/6)(1) = 1/24$ 或大约 4%。注意即使 Jane 的携带者试验结果为阴性,由于她具有较高的前概率(2/3),而且因为该试验的敏感性

在 100%以下,她是携带者的风险仍然有 17%。

因为 Jane 的双亲一定是囊性纤维化杂合性携带者,所以她的一级堂姐妹有 1/4 的机会也是携带者。由于从 Dick 的 DNA 试验中已知他是携带者。Dick 和 Laura 的任何一个子女风险为 1/4(1/4)(1) + 1/16。显然 Laura 的风险也可用 DNA 试验来限定。

5. 这个问题可用 Bayes 定理的条件概率分析或用如表 12.6、12.7、12.8 所示的 2×2 方阵来回答。

假定这名男子无家族史,他是携带者的前概率为群体概率。同时假定 Hardy-Weinberg 平衡,如果该疾病的发病率为 1/10 000,突变等位基因(将所有的突变等位基因看作一个)的频率为 1/100,杂合性携带者的概率为 2×(1/100)×(99/100) = 198/10 000,大约为 1/50 或 2%。有患病同胞的妇女是携带者的前概率为 2/3 。回忆一下,98%的敏感性意味着真正的携带者有 98%的阳性结果,而有 2%的携带者其结果为阴性。90%的特异性表示非携带者 90%为携带者阴性结果,而 10%为阳性结果。

尽管筛查试验有很高的敏感性,因为携带者在一般人群中的发病率十分低,大多数有阳性试验结果的个体可能是假阳性而不是真正的携带者。因此,这名男子实际上不是携带者的概率为 83%。相比之下,当是携带者的前概率高时,同样对于受累个体的非受累同胞来说,绝大多数阳性结果表示是真正的携带者(在下面的例子中 196/206 或 95%)。由于试验敏感性不是 100%,仍有较低的风险(4%)携带者会漏检。

最后,这名携带者试验结果阳性的男子感到困扰不安说明,遗传疾病携带者筛查试验没有相关咨询和教育来辅助绝不能实施。筛查试验需要认真解释,提供准确、有意义的有效风险预测。但是,具有敏感性和辅助性的咨询却尤为重要。疾病的遗传易感性或遗传病传递风险的试验包括重要的情感方面的内容。仅提供这些试验而不提供咨询服务是不恰当的。这些必须考虑的因素在增加分子诊断试验的有用性以预测遗传风险方面特别重要。

Bayes 分析

	男子是携带者	男子是非携带者
前概率	0.02	0.98
后概率(阳性试验)	0.98	0.10
联合概率	0.020	0.098
后概率	$\frac{0.020}{0.020+0.098}=\frac{0.020}{0.118}=0.17$	$\frac{0.098}{0.098+0.020}=0.83$

2×2 方阵,一般人群中杂合子的筛查(筛查了 100 000 人)

	携带者	非携带者	总计
阳性试验	1 960	9 800	11 760
阴性试验	40	88 200	88 000
总计	2 000	98 000	100 000

9 800/11 760 = 83%

Bayes 分析

	女子是携带者	女子是非携带者
前概率	2/3	1/3
后概率(阴性试验)	1/50	9/10 或 45/50
联合概率	2/150	45/150
后概率	$\frac{2/150}{2/150+45/150}=2/47=4\%$	$\frac{45/150}{2/150+45/150}=45/47=96\%$

2×2 方阵,患者同胞中杂合性携带者筛查(筛查了 300 名同胞)

	携带者	非携带者	总计
阳性试验	196	10	206
阴性试验	4	90	94
总计	200	100	300

4/94 = 4%

最后,这名筛查试验阳性结果的携带者所感到的困扰不安说明没有相关的咨询和教育来辅助进行的遗传病携带者筛查试验是不能实施的。对筛查实验需要认真解释,提供准确、有意义、明确的风险预测。具有敏感性和辅助性的咨询就显得极为重要。疾病的遗传易感性实验或疾病遗传的风险预测包括重要的情感方面的内容,仅做这些试验而不提供咨询服务是不适当的。这些必须考虑的因素对于越来越多的分子诊断实验估计遗传病风险是特别重要的。

第十三章 问题

*1. 简要对比甲型血友病和肌营养不良基因治疗的可行性,你认为在今后的十年中那一个更可行? 为什么?

*2. 现在已经通过肺内吸入表达人类 CFTR 的复制缺陷腺病毒的喷雾剂治疗了几例囊性纤维化患者,什么问题将会限制这种方法的疗效? 你认为将会有什么样的替代方法?

3. 反转录病毒
   腺病毒
   腺病毒相关病毒
   非病毒载体
   将上述基因治疗载体与下列特点搭配起来:
   a. 无感染风险
   b. 靶细胞范围广泛, 不需要分裂的细胞
   c. 可能产生插入突变
   d. 整合到染色体特定位点的能力

4. 意外的生殖系基因治疗
   体细胞基因治疗

意向性基因治疗

将上述名词与下列描述相搭配：

a.应用指导 CFTR 表达的腺病毒载体喷雾剂治疗患者，则该患者在治疗后 1 年和 3 年所生的孩子均为 CF 的携带者。

b.ADA 缺陷患者的骨髓细胞在体外经表达 ADA 的反转录病毒处理后进行自体骨髓细胞移植，PCR 分析治疗 1 年后的精子 DNA，检测到反转录病毒的 DNA 序列。

c.应用靶向同源重组的方法进行矫正点突变体外受精的合子。

第十三章 答案

*1.这两种疾病是目前各种基因治疗方法的目标。它们的基因都相当大，所以将面临额外的技术问题，但是肌营养不良基因座尤其困难(抗肌萎缩蛋白大约 450kDa，其mRNA为 14kb，而Ⅷ因子约为 220kDa，其 mRNA 为 9kb)。为了治疗肌营养不良，必须实现基因在肌肉内的特异性高水平表达，因此需要肌肉细胞的高效转导。

与之相反，甲型血友病只需将其蛋白产物分泌入血浆内，因此它可能来自各种不同的细胞类型，包括成纤维细胞、肝细胞、肌细胞或内皮细胞。此外，基因表达的精确调控根本不重要，人体可以耐受较大范围的Ⅷ因子表达水平。甚至表达水平非常低时(1ng/ml)也能发挥临床作用。综上所述，甲型血友病的基因治疗尽管也很困难，但相对于肌营养不良似乎比较容易，在今后的几年内更有望成功。

*2.显然参与这些病例的研究者和患者确信成功的可能性很大，尽管仍然存在着一些潜在性的问题，这些可能限制疗效的潜在性的问题包括感染持续时间往往是有限的，因为 DNA 不能稳定地插入基因组中，而是以附加体的形式存在于细胞中。因此，它会很快地从分裂细胞中丢失。此外，感染细胞可能最终会被免疫系统清除。免疫排斥反应除了能够减少表达细胞，还能造成直接的毒性反应。另外，免疫反应可能会妨碍再次应用同一载体治疗同一患者。尽管 CFTR 能够以这种方式有效地表达，但能够插入现有的腺病毒载体的片段大小受到限制。最后，CF 的病理反应不仅仅限制于肺部，其他组织的病变则不能以此种治疗手段有效地治愈。然而，肺是大多数患者主要的受累器官。替代性方法包括应用反转录病毒载体经体内或转体内、通过喷雾剂使 DNA 直接弥散到肺，或与其他试剂配伍使其更容易进入细胞内部(如脂质体)。

3.

a.非病毒载体

b.腺病毒

c.反转录病毒

d.腺病毒相关病毒

4.

a.体细胞基因治疗。通过喷雾剂将腺病毒载体特异地导入到肺，即使有很少的颗粒被带到生殖腺也不会发生腺病毒的稳定整合，因此，不可能发生载体序列的生殖系传递。因为囊性纤维化是一种常染色体隐性遗传病，患者都为 *CFTR* 基因某一突变

的纯合子(或复合性杂合子)。因此,他所有的子女必定为杂合子。

b.意外的生殖系基因治疗。该实验是按照体细胞基因治疗来设计的,然而 PCR 的结果显示载体序列意外地到达了生殖腺并整合入能够产生精子的生殖细胞。

c.意向性生殖系基因治疗。如第十三章所讨论的,将遗传性改变导入人类生殖系,目前仍不被大多数科学家和医生所接受,但是,此处所描述的过程在转基因鼠实验中属于常规操作。

第十四章 问题

*1.参考病例 1(图 14-1),如果采取一种对 *BRCA1* 突变携带者高度有效、风险又低的治疗策略,对妹妹的启发式方法会如何改变?

*2.在病例 4 中(图 14-4),假定这对夫妇直截了当要求终止妊娠,认为他们不能抚养第五个孩子。然而,与你谈话后,他们认识到羊膜穿刺术可以了解胎儿的性别。假如不能利用产前诊断,他们会选择终止妊娠。如果允许他们了解胎儿性别,他们将允许男性胎儿的妊娠继续,而终止女性胎儿的妊娠。这会改变你提供产前诊断检测的愿望吗?

*3.提供咨询前,你认为下列哪一种情况下应该公开?
出生前或选择前有说服力的个人观点;
公司出资提供患者正在考虑的检测;
对疾病的研究关注;
根据以前咨询过其他家系成员,个人熟悉系谱方面的情况。

*4.一位对 *BRCA1* 感兴趣的病理医师已经接触了大量乳腺癌标本,这些标本来自医院外科病理实验室。利用 PCR 可能扩增石蜡包埋的 DNA 进而识别基因突变。他研究了 25 例乳腺癌,发现其中一例存在 *BRCA1* 突变。分析包埋中的癌旁正常组织,发现同样的 *BRCA1* 突变,表明此突变存在于生殖细胞。关注到这位妇女可能有乳腺癌或卵巢癌的附加风险,他要来她的病志,然后给她打电话告诉他的发现。期望因他的努力而得到感谢,令他吃惊的是她反而质问没有她的许可他怎么能这样做。这里有什么违反伦理之处?

第十四章 答案

*1.当有几种未经证实益处的选择时(如本例中的 *BRCA1* 突变携带者),医务人员的启发式态度尤其重要。在那种情况下,尊重患者自主权会削弱权威性的"医生知道怎样最好"的方案。但是如果治疗显示有明显的益处,那么行善原则要求医务人员态度更有指导性。例如,儿科医生面对小儿急性脑膜炎病例时,应明确建议使用抗菌素。

*2.满怀希望的父母的表现的微妙变化会影响医务人员的决策。如果他们确信没有羊膜穿刺术肯定会终止妊娠,他们可能会默许对羊膜穿刺术的要求。

*3.当医务人员不可能是完全客观的时候,公开揭示利益冲突值得深思。如果这些冲突对病例并不重要,患者了解这些也会是痛苦的。在一些例子中(即研究利益),潜在的

利益冲突实际上可能是带来益处。遗传咨询员采取的启发式态度从理论上讲可以防止由于咨询顾问的偏见而改变结果——尽管不经常达到这种水准。当医务人员意识到潜在的利益冲突会影响他(她)对病例的判断时,和一个客观的同事探讨可能是最明智的选择。如果情况竟是咨询员不能公正地起作用,则应表明利益冲突,换一个咨询员来咨询。

*4.在没有知情同意的情况下,病理医师对一个可确认来源的标本进行研究,这样做侵犯了患者的自主权(她的知情权利被践踏)和她的隐私。大多数研究者认为,进行研究前,仅在标本的个人身分标签不慎丢失而不能确定标本来源时,才可以进行这种未经同意的研究。然而,如果个人身分可以识别,即使标本被编号,实际上经常需要患者的知情同意。

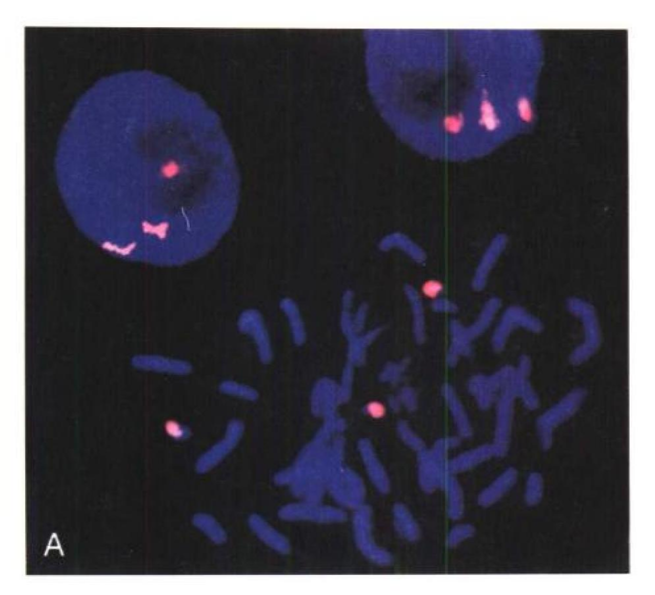


图 8-8 FISH 的范例

A. 使用一个21号染色体涂染探针，应用于21号染色体三体性Down综合征患者的一个中期和两个间期细胞的FISH分析。在中期细胞中，21号染色体长臂全长被涂染。涂染探针不包括21号染色体短臂的序列，因为这一区域的探针会和其他近端着丝粒染色体发生交叉杂交。在间期细胞，尽管单个染色体无法鉴别，但三个清晰的“显像”是明显的（引自 DL Van Dyke and A Wiktor,Henry Ford Health Sciences Center.）。

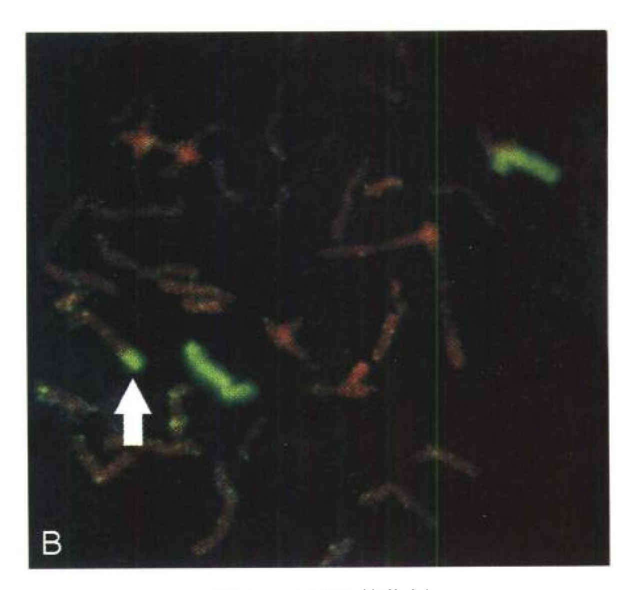


图 8-8 FISH 的范例

B. 携带10号染色体不平衡常染色体易位的患者中期染色体的FISH分析。通过标准的Giemsa染色，10号染色体上多余物质的来源不能确定. 应用9号染色体涂染探针，两条正常的9号染色体沿其全长被涂色，而异常的10号染色体（箭头）显示含有一个易位到10号染色体长臂的9号染色体物质的重复片段（引自 DL Van Dyke and A Wiktor,Henry Ford Health Sciences Center.）。

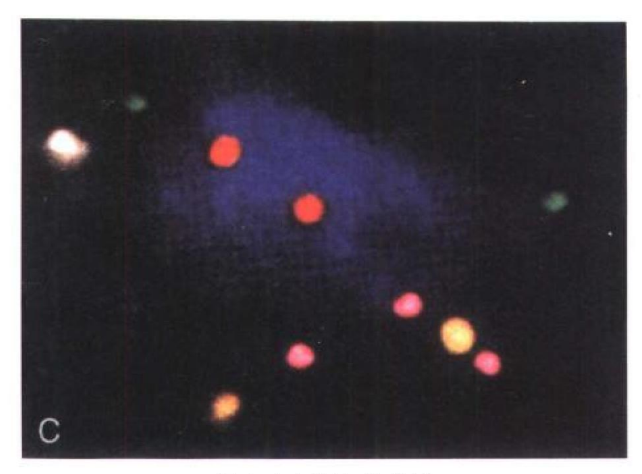


图 8-8 FISH 的范例

C. 应用 5 个着丝粒探针的一个间期细胞的多色 FISH（未培养的羊水细胞）。13 号染色体（红色），21 号染色体（绿色）各含有两个信号，但 18 号染色体信号（粉色）有三个信号，还有两个 X 染色体信号（黄色）和一个 Y 染色体信号（白色）。这样，该细胞显示为 18 号染色体和性染色体的三体性。核型是 48，XXY,+18（引自 Ried T, Landes G, Dackowski W, Klinger K, Ward DC: Multicolor fluorescence *in situ* hybridization for the simultaneous detection of probe sets for chromosomes 13,18,21,X and Y in uncultured amniotic fluid cells.Hum Mol Genet 1992;1:307~313.）。

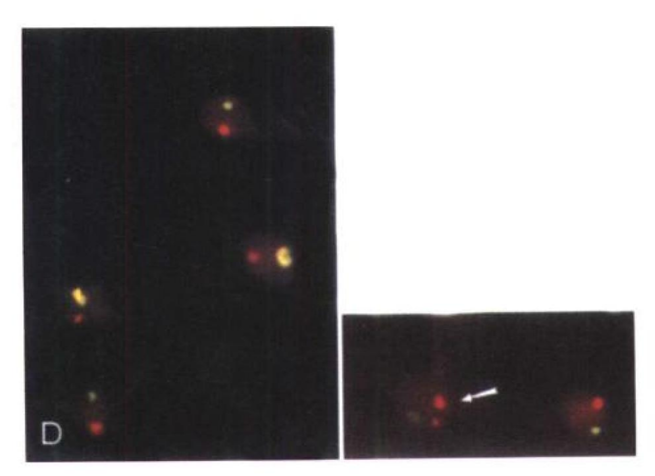


图 8-8 FISH 的范例

D. 人精子染色体不分离的 FISH 分析。应用 8 号染色体（红色），Y 染色体（黄色）和 X 染色体（绿色）着丝粒探针三色杂交。在左侧显示四个正常的单倍体精子，而在右侧显示一个 8 号染色体二体性精子（箭头）（引自 Williams BJ, Ballenger CA, Malter HE, Bishop F, Tucker M, Zwingman TA, Hassold TJ: Non-disjunction in human sperm: results of fluorescence *in situ* hybridization studies using two and three probes. Hun Mol Genet 1993;2:1929~1936.）。

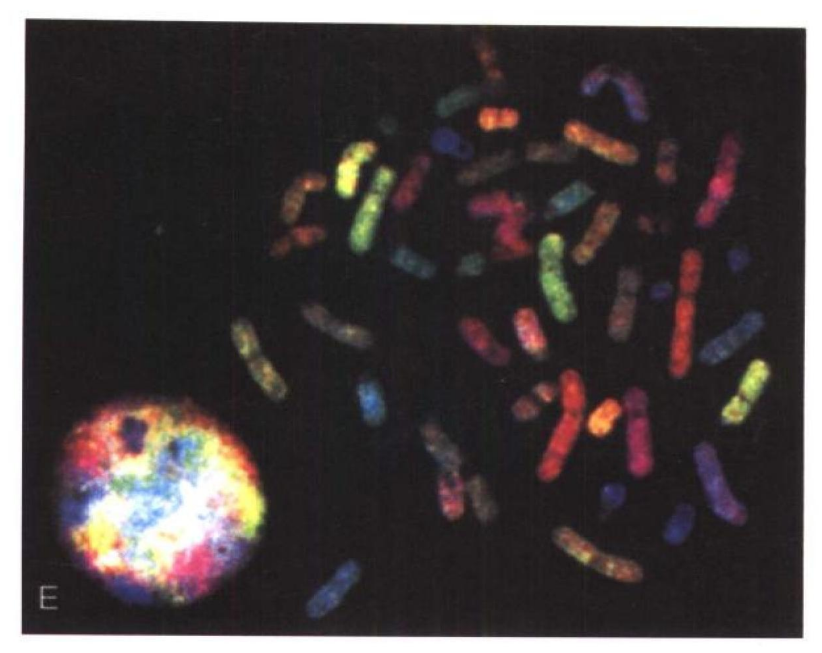


图 8-8 FISH 的范例

E. 一个正常人中期涂片与 5 种不同荧光素混合标记的 24 种不同的染色体涂染探针杂交后的色谱图像（引自 T Ried, National Center for Human Genome Research.）。



图 8-8 FISH 的范例

F. 应用 24 色多重 FISH（M-FISH）的核型，该核型来自携带 2 号染色体短臂和 14 号染色体长臂相互易位的患者（引自 Speicher MR, Ballard SW, Ward DC: Karyotyping human chromosomes by combinatorial multi-fluor FISH.Nat Genet 1996;12:368~375.）。

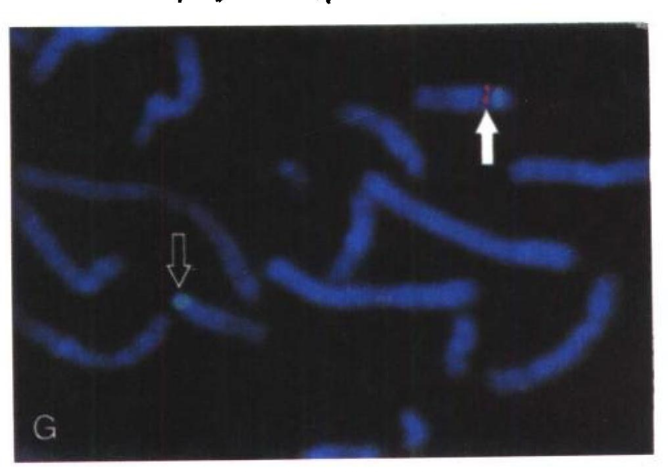


图 8-8 FISH 的范例

G. 应用基因座特异性黏粒探针（红色，实心箭头）对Prader-Willi综合征关键缺失区域内臂间缺失的FISH诊断。用一个15号染色体特异性着丝粒探针（绿色）鉴定了两个15号同源染色体。有缺失的一条同源染色体（空心箭头）显示与PWS区域黏粒探针无杂交（引自 D Ledbetter, National Center for Human Genome Research.）。